The application of quantitative trade-off analysis to guide efficient medicine use within the UK NHS

by

Pritesh Naresh Bodalia

MPharm., PgCertPP., PgDipGPP., MSc., MRPharmS.

A thesis submitted for the award of PhD

Faculty of Population Health Sciences

Institute of Epidemiology and Health Care

University College London

March 2018

I, Pritesh Naresh Bodalia confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Background: The UK National Health Service (NHS) currently spends in excess of £17 billion per annum on medicines. To ensure sustainability, medicines with limited clinical value should not be routinely prescribed. The randomised controlled trial (RCT) aims to meet regulatory standards, therefore it has limited value to guide practice when multiple agents from the same therapeutic class are available. The quantitative trade-off analysis, with consideration to efficacy, safety, tolerability and / or cost, may enable the translation of regulatory data to clinically-relevant data for new and existing medicines.

Methods: This concept aims to provide clarity to clinicians and guideline developers on the efficient use of new medicines where multiple options exist for the same licensed indication. Research projects of clinical relevance were identified from an academically led London Area Prescribing Committee (APC). Therapeutic areas included refractory epilepsy, hypertension and heart failure, overactive bladder syndrome, and atrial fibrillation. Frequentist and / or Bayesian meta-analysis were performed and supplemented with a trade-off analysis with parameters determined in consultation with field experts.

Results: The trade-off analysis was able to establish a rank order of treatments considered thereby providing clarification to decision-makers on the DTC / APC panel where regulatory data could not. The results, presented as a hierarchy of treatments, enabled modifications to prescribing trends within North Central London as the pilot site, resulting in significant cost avoidance and cost savings for the NHS.

Conclusions: The quantitative trade-off analysis was able to resolve concerns raised by the DTC / APC panel via translation of regulatory data to clinically-relevant data with consideration to defined benefits and harms. Results were implemented successfully within a local pilot health economy. This approach is recommended as an extension of existing methods required by regulatory agencies in the assessment and licensing of new medicines.

Impact Statement

The UK National Health Service (NHS) currently spends 15% of its £116 billion budget per annum on medicines; this is in excess of £17 billion, having grown from £13 billion in 2011-12. To ensure that the most effective medicines are available to patients, provider Acute Trusts have established local and / or regional Drugs and Therapeutics Committees (DTCs) to maintain a local formulary. Despite being the gold standard in evidence based medicine, randomised controlled trials (RCT) are designed to meet regulatory standards; therefore it has limited applicability in guiding clinical practice, particularly when multiple agents from the same therapeutic class are available. Guidance published by the National Institute for Health and Clinical Excellence (NICE) determines the clinical and cost-effectiveness of each individual new medicine, where the output is either the recommendation or non-recommendation of said medicine; where it is recommended it has to be made available for prescribing among other recommendations. Establishing the place in therapy for new and / or existing treatment as part of the DTC review is therefore often unclear.

The UCL Research Department of Epidemiology & Public Health aim to provide leading and multi-disciplinary research with impact on real-world policy and practice. The quantitative trade-off analysis is a concept that gives consideration to key efficacy *versus* acceptability *versus* cost parameters for a range of similar medicines licensed for a given therapeutic indication, and aims to provide clarity to clinicians and guideline developers where multiple options exist. The methodology employed in order to undertake this analysis is based on Bayesian statistics, using network meta-analysis (NMA) techniques, via the WinBUGS software. Specific therapeutics areas explored include: refractory epilepsy; cardiovascular disease (hypertension and heart failure); overactive bladder syndrome; and atrial fibrillation.

The research themes and projects were conceived in conjunction with a multi-disciplinary Area Prescribing Committee (regional DTC) within London and made possible via effective collaboration across a number of academic organisations including University College London, UCL Hospitals, University of Bristol, and the London School of Hygiene and Tropical Medicine. Specific areas of collaboration included individuals with a background in statistics

modelling and methodology, clinical / medical, health economics, and patient partnerships.

The application of the quantitative trade-off analysis has successfully aided in the establishment of a hierarchy of treatments where regulatory data could not, and has been used to change practice across a pilot area of the UK NHS resulting in cost savings or cost-avoidance without compromising on patient care. Indirect comparisons, generated through NMA as conducted across a number of clinical specialities, not only helped develop rational treatment hierarchies, but also provide guidance on the choice of high priority comparator agents for direct head-to-head analysis in the form of an RCT. This approach is recommended as an extension of existing methods required by regulatory agencies in the assessment and licensing of new medicines. The outputs of the analyses conducted have been published in leading medical journals to inform wider practices.

Acknowledgements

I am grateful to Prof Aroon Hingorani and Prof Juan-Pablo Casas for their supervision, time and input. Similarly, Prof Raymond MacAllister and Dr Anthony Grosso have been invaluable in their assistance and enthusiasm throughout. I am also particularly grateful to Dr Robert Urquhart and Mr Ashik Shah for their continued mentorship for my research and development activities.

I would also like to thank Mr David Wonderling for his advice and tuition on the use of WinBUGS in order to perform the Bayesian Network Meta-Analyses within this thesis and Dr Nicola Welton for her advice and support on developing techniques to the standard of the NICE [Technical] Decision Support Unit. In addition, I am indebted to Dr Michael Scott and the health economics team at University of Bristol for their contribution to the thesis in performing cost-utility analyses. I offer thanks to Prof Jonathan Sterne for leading the NIHR project which forms a section of this thesis, and to my teams, both within UCL Hospitals Pharmacy and from the NCL Joint Formulary Committee, for their support and understanding.

I am thankful to the numerous experienced academics who have given time to perceptively critique this research, including the unknown peer-reviewers selected by the medical journal editors.

Finally, I am, as ever, grateful to my family, wife and two boys for their continued encouragement without exception.

Table of Contents

List of Abbreviations

AAN	American Academy of Neurology
ACEI	angiotensin-converting enzyme inhibitor
AE	adverse effect / adverse event
AED	antiepileptic drug
AES	American Epilepsy Society
AF	atrial fibrillation
AIIRA	angiotensin-II receptor antagonist
APC	Area Prescribing Committee
bd	twice daily
BMI	Body mass index
BNF	British National Formulary
BP	blood pressure
BUGS	Bayesian analysis using Gibbs Sampling
CADTH	Canadian Agency for Drugs and Technologies in Health
CCG	Clinical Commissioning Group
CCTR	Cochrane controlled trials register
CEA	Cost effectiveness analysis
CEAC	Cost-effectiveness acceptability curve
CER	comparative effectiveness research
$CHADS_2$	a clinical prediction rule for estimating the risk of stroke in patients with non-valvular atrial fibrillation
$CHADS_2VASC$	a modified clinical prediction rule for estimating the risk of stroke in patients with non-valvular atrial fibrillation
CHD	coronary heart disease
CHF	congestive heart failure
CI	confidence intervals (frequentist analysis)
CMA	cost minimisation analysis
CPRD	Clinical Practice Research Datalink
CRB	clinically relevant bleeding
CRD	Centre for Reviews and Dissemination
CrI	credible intervals (Bayesian analysis)
CRNM bleeding	clinically relevant non-major bleeding
CUA	cost utility analysis
CVD	cardiovascular disease
DBP	diastolic blood pressure
DIC	deviance information criterion
DSU	Decision Support Unit
DTC	Drugs & Therapeutics Committee
ECG	electrocardiogram
EMA	European Medicines Agency
FAS	full analysis set
FDA	Food and Drug Administration
FE	fixed-effect

HAS-BLED	hypertension, abnormal renal/liver function, stroke, bleeding history or predisposition, labile international normalized ratio, elderly (> 65 years), drugs/alcohol concomitantly: this estimates risk of major bleeding for patients on anticoagulation for atrial fibrillation
HR	hazard ratio
HTA	Health Technology Assessment
ICER	incremental cost-effectiveness ratio
ICH	intracranial haemorrhage
ICS	International Continence Society
ILAE	International League Against Epilepsy
INB	incremental net benefit
INR	international normalised ratio
IPD	individual patient data
ITT	intention to treat
LMWH	low molecular weight heparin
LOE	languages other than English
MCe	Monte Carlo error
MCMC	Markov Chain Monte Carlo
MD	mean difference
MHRA	Medicines and Healthcare Regulatory Agency
MI	myocardial infarction
MTC	mixed treatment comparison
NCL	North Central London
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NIHR	National Institute for Health Research
NMA	network meta-analysis
NMB	net monetary benefit
NNH	number needed to harm
NNT	number needed to treat
NOAC	novel oral anticoagulant
NR	not reported
NVAF	non-valvular atrial fibrillation
OAB	overactive bladder
od	once daily
ONS	Office for National Statistics
OR	odds ratio
PP	per protocol
PRISMA	preferred reporting items for systematic reviews and meta-analyses
QALY	quality adjusted life year
QUOROM	quality of reporting meta-analyses
RAAS	renin angiotensin aldosterone system
RCT	randomised controlled trial
RE	random-effects
RRR	relative risk reduction
SBP	systolic blood pressure

SD	standard deviation
SE	systemic embolism
SIGN	Scottish Intercollegiate Guidelines Network
SMC	Scottish Medicines Consortium
SMD	standardised mean difference
SUCRA	surface under the cumulative ranking curve
TA / TAG	Technology Appraisal (Guidance)
TDM	therapeutic dose monitoring
TIA	transient ischaemic attack
TTR	time in therapeutic range
UUI	urgency urinary incontinence
VKA	vitamin K antagonist
VTE	venous thromboembolism
WinBUGS	a Microsoft operating system version of BUGS (Bayesian analysis using Gibbs Sampling)
WMD	weighted mean difference

List of Figures

List of Tables

1.0 Chapter 1: Introduction

1.1 Introductory Statement

Rational prescribing of medicines has been the cornerstone of modern medicine and is increasingly an important area for research for many reasons. The UK NHS currently spends 15% of its £116 billion budget on medicines. This has increased from £13 billion in 2011/12 to £16.8 billion in 2015/16.(1) Spend in London (£2,852.8m) is the highest for all NHS England regions owing to its higher population.(2) The regulatory process, undertaken by the European Medicines Agency (EMA) is well established in the assessment of the clinical effectiveness of new medicines. The Technology Appraisal process undertaken by the National Institute for Health and Clinical Excellence (NICE) goes one step further to determine the clinical and cost-effectiveness of each individual new medicine. However, guidance offered from such appraisals is either the recommendation or non-recommendation of the new medicine; where if recommended it has to be made available for prescribing among other recommendations.

The current standards for approval or recommendation therefore fail to assess whether new treatments are better or less efficacious or tolerated than existing alternatives. The place in therapy for the new and / or existing treatment therefore is often unclear and becomes compounded with every additional new medicine introduced into the market. With effective marketing techniques this introduces the possibility that patients may be harmed or less effectively managed through receiving newly approved treatments instead of alternative established treatments. In addition, the UK health economy also suffers as new treatments prescribed on the NHS are considerably more expensive, therefore providing these requires a disinvestment from other areas. A clear understanding of the potential benefits (efficacy) and harms (safety / tolerability) is needed to determine this.(3)

Comparative Effectiveness Research (CER) has gained interest by the US FDA [Food and Drug Administration] as part of their process for approval of a new medicine. Although the European Medicines Agency

(EMA) has supported the requirement of active-comparator studies since 2004, the full spectrum of CER remains a new concept.

The quantitative trade-off analysis is used in many areas of marketing new products to consumers, but remains a new concept to healthcare, in particular for stratification of place in therapy for new medicines. The submission of trade-off analysis as part of the dataset required by regulatory authorities prior to the approval and widespread adoption of a new medicine has the potential to reduce unnecessary spend on inferior or more costly equivalent treatments.

This thesis explores the quantitative trade-off analysis across four therapeutic areas, namely refractory epilepsy, cardiovascular disease (hypertension and heart failure), overactive bladder syndrome, and atrial fibrillation. A brief introduction into their aetiology, epidemiology, clinical manifestations and treatment is provided within each section. Pertinent clinical information concerning key efficacy and adverse events are also described as part of the appraisals which are used to develop the trade-off analyses.

1.2 Drug Appraisal Process in Clinical Trials

Clinical Trials are research studies involving human subjects, often patients, which aim to evaluate the safety and effectiveness of a new treatment. Clinical trials are the most reliable method of evaluating new treatments provided their design objectives have been appropriately met.

1.2.1 Phase I clinical trials

In phase I clinical trials the investigational drug is given to humans for the first-time. These studies are usually conducted in healthy volunteers. However, there are some circumstances when patients are used, often in situations where patients may have a severe or rare disease where there are a lack of adequate treatment options currently available thereby constituting an unmet clinical need. The emphasis of the design of phase I studies is to determine the metabolic and pharmacological actions of the investigational drug in humans, including any adverse events associated with incremental

dose increases, and, if possible, to gain any early evidence on clinical effectiveness. The total number of subjects included in Phase I studies is often small, usually in the range of 20-80.

1.2.2 Phase II clinical trials

Phase II clinical trials are the early controlled clinical studies conducted to obtain some preliminary data on the effectiveness of the drug for a particular indication [or indications] in patients with the actual disease or condition. This phase of testing also helps determine the common short-term adverse effects and risks associated with using the drug in the proposed patient population. Phase II studies usually involve a few hundred patients and such data will be used to inform on the dosage and design of the larger, usually multi-centre, phase III studies. Many adverse effects are still unlikely to be detected due to the small patient numbers. The observed efficacy may still be a chance finding as the power of the study will be low.

1.2.3 Phase III clinical trials

Phase III clinical trials are intended to gather pivotal clinical information about the effectiveness and safety of a new treatment that will be used to evaluate its overall risk:benefit relationship. Phase III trials involve larger numbers of patients (hundreds or thousands) who are usually randomised to receive the new treatment or the current best available (including placebo). The aim is to assess how well the new treatment works in a meaningful number of patients which meet a statistical power in order to extrapolate or generalise the findings from the sample population to the general population. The larger numbers also supports identification of serious adverse effects, however, the rigid design often precludes the recruitment of ‘at-risk’ patient groups. The aim of the phase III clinical trial is to satisfy regulatory criteria in order to obtain a Marketing Authorisation.

1.2.4 Phase IV clinical trials

Phase IV clinical trials are conducted to identify and evaluate the long-term effects of new medicines and treatments over a lengthy

period for a greater number of patients. Phase IV research occurs following regulatory approval and marketing of a new medicine, with data capture primarily from the primary care (general practice) setting or routine specialist care. Although the protocol for such pharmacovigilence plans may include assessment for efficacy, safety and side-effects in the post-marketing period, data collection is largely centred around safety and tolerability.

1.3 Pharmaceutical Industry

Clinical research sponsored by the pharmaceutical industry influences medical practice because the vast majority of clinical trials at all stages in a product's life cycle are funded by the pharmaceutical industry. Results that are unfavourable to the sponsor - that is, trials that find a new medicine is less clinically-effective or cost-effective or even more harmful than other medicines used to treat the same condition, will thus pose considerable financial risks to companies. Pressure to show that the medicine causes a favourable outcome may result in biases in design, outcome, and reporting of industry-sponsored research.(4) Examples of such potential biases are outlined below and in section 3.4.6.

1.3.1 Design Bias

Funding may promote study designs that are more likely to produce favourable results, such as designs involving only placebo or other poor comparators, inappropriate doses, carefully constructed experimental populations, poor surrogate endpoints, trial durations unlikely to show adverse effects, and protocols likely to show activity or unlikely to show adverse effects. An example of such biased design was seen in the trial of a new proton-pump inhibitor where the Sponsor compared the new medicinal product, esomeprazole (an active isomer of the racemic comparator, omeprazole), to a clinically inappropriately low dose of omeprazole.(5)

1.3.2 Data Interpretation

Industry-supported trials and reviews of drugs should always be read with caution as they have been shown to be less transparent, have fewer reservations about methodological limitations of the included trials, and have more favourable conclusions compared to independent reviews.(6) A recent study published in the *British Medical Journal* reviewed 24 meta-analyses; eight were industry supported, nine had undeclared support, and seven had no support or were supported by non-industry sources. Compared with industry-supported reviews and reviews with undeclared support, independent reviews had more often considered the potential for bias in the review; for example, they described the method of concealment of allocation, excluded patients or excluded studies. The seven industry-supported reviews that had conclusions recommended the experimental drug without reservations, compared with none of the independent reviews. 'Ghost-writing' is also an issue as reports can be researched and written by, or on behalf of, pharmaceutical companies, and then published under the name of academics who had played little role in the research and writing process. The resulting articles affect the conclusions found in the medical literature, and are used in promoting drugs to healthcare providers.(7)

For industry-funded trials, positive data are over-reported relative to negative data.(8) There may be other kinds of publication bias as well. Almost all journals earn considerable revenue from sales of reprints of articles sometimes selling hundreds of thousands of articles with high commercial value. This creates potential conflicts of interest that could affect publication patterns. Many journals earn money from the publication of supplements, often based on symposia sponsored by pharmaceutical companies. Peer-review of these supplements typically differs from those of normal issues of these journals. A recent analysis of new drugs approved by the FDA between 1998 and 2000 demonstrated that selective reporting of trial results occurred for commonly marketed drugs.(8) Moreover, over one half of all supporting trials for FDA-approved drugs remain unpublished at five years [or greater] post-approval.(9)

The marketing budgets of the drug industry are enormous although exact figures are difficult to come by. This is in part because marketing and administrative expenses are often grouped together and in part because some of the research and development budget is for marketing research.(10) Many healthcare professionals rely on drug company representatives and promotional materials to learn about new drugs. In the US, the pharmaceutical industry also has direct access to the general public where direct-to-consumer advertising is permitted [such advertising is not allowed for prescription-only medication in the UK]. In summary, one should remain vigilant that the commercial imperative is likely to result in an inappropriate influence on presentation of drug data.

1.4 Regulatory Agencies

Following the analysis of pre-clinical and clinical data, the relevant pharmaceutical company is required to apply for a Marketing Authorisation [product licence] and receive approval before the new medicine can be marketed and prescribed. This is done through the European Medicines Agency (EMA) for all European countries including Norway and Iceland [centralised route] or via the Medicines and Healthcare products Regulatory Agency (MHRA) in the UK [national decentralised route].(11)

Where the regulatory agency concludes that the new drug's benefit profile exceeds its risk profile, and presents an improvement in therapy in relation to its trial comparator, a Marketing Authorisation will be granted. However, as regulatory agencies typically only require evidence of effectiveness over placebo, this presents an issue for clinicians and commissioners of medicines in determining which new treatment should be made available from the range of options currently in the market. Since 2004 the EMA have mandated the requirement of an active comparator where a licensed medicine exists, however, this only provides data on one alternative where in practice a number are available. Furthermore, this process is not able to account for treatments that may be developed in tandem and therefore the use of a placebo or a historic control will be used. As such, the benefit and

harm profile of new medicines as compared with interventions already established in clinical practice are unknown.

1.5 Commissioning Agencies

Expenditure on pharmaceuticals is the fastest growing sector within healthcare in developed countries, including Canada,(12, 13) the United Kingdom,(14) Australia,(15) and the United States.(16, 17) Cost [and cost-effectiveness] is therefore an important consideration which has driven the recent proliferation of 'fourth hurdle' systems. In the UK, NICE was established in April 1999 which considers select new drugs for an evidence-based clinical and cost-effectiveness assessment to determine whether such therapies should be listed in public formularies. In Scotland, a similar body known as the Scottish Medicines Consortium [SMC] evaluates all new drugs before they can be prescribed within NHS Scotland. In London, a body akin to the SMC is currently under development. Although valuable, the outputs of these systems remain restricted to independent recommendations rather than a holistic assessment of all treatments available and therefore the corresponding place for each of these.

1.6 Drug Evaluation in Clinical Practice

In addition to regulatory authorities, such as the EMA, and national commissioning organisations, such as NICE, most local health economies exert local control over what drugs should be made available to prescribers [and subsequently to patients]. This is achieved through the management of local formularies where decisions on formulary entry and restrictions are made by a local drugs and therapeutics committee [or equivalent]. These committees evaluate the comparable efficacy and safety profiles of new medicines versus other available treatment options when deciding on the potential advantages and disadvantages of formulary inclusion; however this is done to varying degrees dependent upon local expertise. Additional considerations are given to convenience and cost.

For example, a newly launched, EMA licensed, beta-adrenoceptor blocking drug is unlikely to gain approval for inclusion to a hospital

formulary if there are no robust data indicating that it is more effective than existing [often cheaper] options. Many new drugs do have some potential advantages, for example, once daily dosing in comparison to a twice daily regime, however such advantages may be considered insufficient to offset increased drug costs or lack of longer term safety assurances. Further, trials reporting statistical significance may not be indicative of clinical significance i.e. a difference that will make a meaningful difference to a patient and the treatment or management of their condition. Decisions from such committees across the UK are variable and are often associated with a high rate of approval. Per capita wastage of this type tends to be greatest in hospitals; this could be reduced if some simple principle of drug management and use were followed.

Medicines evaluation of this type is highly developed at University College London Hospitals, which is coordinated by the Use of Medicines Committee (UMC), as identified by the House of Commons Select Committee on Health as part of their review on control of access to medicines.(18) This team of Pharmacists and Clinical Pharmacologists are closely involved in the business of the NCL Area Prescribing Committee.

1.6.1 Spend on Medicines within the NHS

Over the period 2016/17, the cost of NHS medicines prescribed in hospitals and the community totalled £17.4bn.(1) As expected, the highest area of spend is in London at £2,852.8m. This represents a 34% rise since 2011/12 where the annual spend was £13.0bn. Spend in the community setting (primary care) accounted for 56.1%. With an ageing population, newer medicines coming to market with a larger price tag, and a finite NHS budget, it is important that medicines are prescribed wisely to offer as much as possible to the population as a whole.

1.6.2 Clinicians and Allied Healthcare Professionals are close to the research ideas

The best health research ideas, and certainly those that are most applicable, come from clinical practice, especially through working with patients and the public. This is precisely the setting where Pharmacy Practice Research excels and underpins this thesis. Medicines Management teams based within an NHS hospital setting undertake clinical care for patients, both as individuals and at service level. Through discussions with colleagues, a number of topics of research were developed to form the chapters of this thesis as part of enhanced comparative effectiveness research.

1.7 Comparative-Effectiveness Research

Comparative effectiveness research (CER) is a relatively new concept which is designed to inform health-care decisions by providing evidence on the effectiveness, benefits and harm of different treatment options.(19) The evidence is generated from research studies that compare drugs, medical devices, tests, surgeries, or ways to deliver health care.

There are two ways that this evidence may be found:

- Researchers look at all of the available evidence regarding the benefits and harms of each drug available for the treatment or management of a specific condition from *existing* clinical trials, clinical studies, and other research. This is more commonly referred to as a *systematic review* of the existing literature.
- Researchers conduct studies that generate *new* evidence of effectiveness or comparative effectiveness.

CER requires the development, expansion, and use of a variety of data sources and methods to conduct timely and relevant research and disseminate the results in a form that is quickly usable by clinicians, patients, policymakers, and other payers.

In general there are six steps involved in the conducting this type of research and ensuring continued development of the infrastructure to sustain and advance these efforts:

1. Identify clinical interventions available
2. Review and synthesise current research
3. Identify gaps between existing medical research and the needs of clinical practice
4. Generate new scientific evidence and analytical tools.*
5. Promote this new evidence by disseminating to medical practitioners and relevant stakeholders
6. Translate the findings into clinical practice

*New evidence may include performing one or more of the following analyses: cost-minimisation, cost-effectiveness, cost-benefit or cost-utility.

The aim of CER is to provide an informed decision, based on the best possible evidence, on what the best treatment(s) is. In the absence of changes regarding the pathway of medicines obtaining a marketing authorisation as enforced by regulatory agencies, CER provides the most comprehensive overview for both the provider and commissioner.

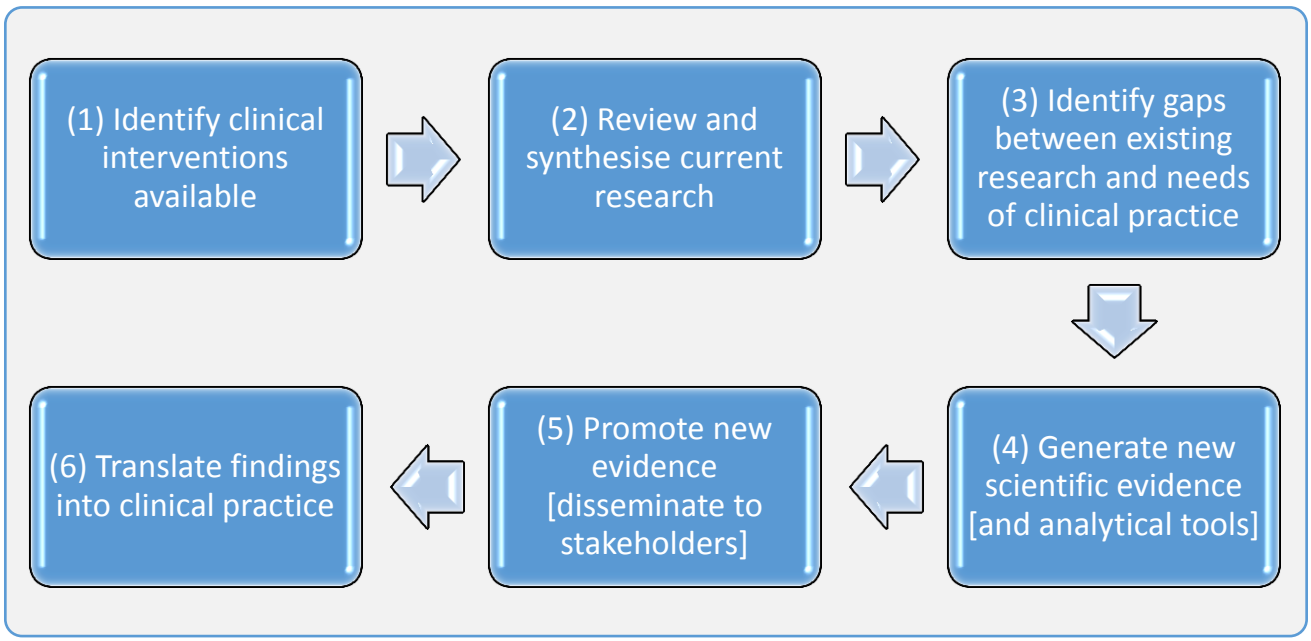


Figure 1: Process of Comparative Effectiveness Research (CER) from concept to practice

1.8 Pharmacy Practice Research

The scope of pharmacy practice research is an umbrella term for research into pharmacy services, medicines use, professional practice and education.(20) New directions in healthcare policy, as well as changing structural, economic, social and cultural contexts of healthcare, and the aspirations of pharmacists for a greater role in the delivery of medicines provide the framework and background for the conception and execution of pharmacy practice research. The ultimate goal is to lead the way in adaptation of services with participation in original research seen as a fundamental component.

Topics commonly addressed under this area of research include:

- Pharmaceutical service provision, delivery and development
- Quality and safety of services
- Therapeutic outcomes consequent on service provision and medicines use
- Pharmacoeconomics, costs and cost-effectiveness
- Assessment of medicines-related needs from the perspectives of patients and carers
- Health and pharmaceutical policy

1.9 Trade-off Analysis

Every pharmaceutical manufacturer of a new medicine approaches launch and marketing of their new product from the business perspective, whilst clinicians and healthcare professionals will view its clinical opportunities. A tool used by the Marketing Research community is the ‘trade-off analysis’ to support marketing decisions that require complex decision making incorporating several data points. Fundamentally, this approach assigns attributes to a particular product that is considered most important to its target consumers. Qualities of the trade-off analysis include:

- An ability to differentiate between products so that recipients in the marketplace gain clear direction for allocating their resources towards consumer benefits
- A flexible model to adapt to the relative importance of a number of factors and variables

- Simplicity in the research process that is transparent, easy to understand and may be replicated
- User-friendly results so that marketing applications are clear and actionable

Within the context of this thesis, the trade-off analysis refers to a balance of defined efficacy and safety / tolerability parameters, as determined in conjunction with experts within the field that the research project relates to. The trade-off methodology would fall under the umbrella of Comparative Effectiveness Research as indicated by step 4 of the process depicted in Figure 1. The results would be presented in table or graphical format to enable ease of transfer into a guideline, where required, for dissemination to relevant stakeholders.

2.0 Chapter 2: Aims & Objectives

2.1 Rationale for the Research Project

The NHS governance structure for the management of medicines provides delegated authority to local and / or regional Drugs and Therapeutics Committees (DTC) as well as local health economy commissioners. Individuals performing these tasks are required to make inferences from trial data when considering the merits of new medicines within the treatment pathway alongside existing therapy.

Although NICE provide recommendations via its Technology Appraisal Guideline (TAG) publications on which therapies confer cost-effectiveness within the UK NHS setting, they do not provide firm recommendations of which medicine(s) should be preferentially prescribed from a series of options within the same class. As such, variability in prescribing and therefore expenditure continues to occur across the UK.

This fundamental approach will be applied to four areas of NHS business, as highlighted by the North Central London Joint Formulary Committee (NCL JFC) with adaptations to methodological techniques as necessary. The topics chosen impact primary care prescribing on a national basis.

2.2 Research Question

"Can the use of quantitative trade-off analysis guide efficient use of medicines within the UK NHS."

2.3 Aims and Objectives

The aim of this research is to establish the use of post-trial quantitative trade-off methodology, applied to data obtained from randomised controlled trials to aid clinicians, medicines management committees, and commissioners in their decision making process to promote appropriate and cost-effective prescribing of medicines within the NHS.

The objective of this research is to demonstrate the value of a quantitative trade-off methodology and the benefits it may realise within the UK NHS.

The impact of this is to allow consideration of improved post-marketing prescribing that could be enforced by policymakers such as medicines management or area prescribing committees. Study specific objectives were as follows:

1. Assessment of the efficacy and tolerability profile of old versus new anti-epileptic drugs for the management of refractory focal epilepsy

2. Assessment of the efficacy, tolerability and cost-effectiveness profile of angiotensin receptor blockers (candesartan and losartan) for the management of hypertension and heart failure

3. Assessment of the efficacy and tolerability profile of old versus new antimuscarinic drugs for the management of overactive bladder syndrome

4. Assessment of the efficacy, tolerability and cost-effectiveness profile of all novel anticoagulation therapies compared with warfarin for the prevention of stroke in AF

The rationale for the choices of the exposures and outcomes selected for the investigations were as follows:

- The medicines investigated are used for important chronic medical conditions where therapy is currently non-curative. Patients will therefore continue to receive these medicines, with new medicines within class continuing to emerge resulting in high spend for the NHS. Moreover, the conditions are relatively common meaning that the potential for exposure is high and hence safety / tolerability concerns as well as avoidable cost could have a substantial impact on a population scale.
- The potential adverse event outcomes under investigation are associated with considerable morbidity and mortality, such as

head injury following uncontrolled seizures and stroke following poorly managed atrial fibrillation.

2.3.1 Peer review assessments

The projects undertaken as part of this thesis have undergone peer review assessment as indicated within their respective chapter. Where published, abstracts have been re-produced in the appendices (Appendix VII to Appendix X).

2.4 Disclosure

2.4.1 Ethics and Scientific Approval

As data required for systematic reviews, meta-analysis, and or cost-utility analysis are derived from publically accessible content, neither ethics nor scientific approval were required.

2.4.2 Funding

All data contained within this thesis was available free of charge as publically accessible content from published articles or as source data directly from the author or institution. A research grant was received from the National Institute for Health Research (NIHR) to facilitate progression of the topic relating to novel anticoagulants. However, the NIHR had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. No academic or commercial funding was obtained for any of the other topics within the thesis.

2.4.3 Collaboration

For areas of conjoint work, the specific areas delivered by me and that undertaken by other parties are described within chapter 4 to chapter 7 under ‘sub-section x.4 collaboration.’

[BLANK PAGE]

3.0 Chapter 3: Methods (Data Management & Statistical Analyses)

3.1 Data acquisition

Advances in information technology have enable ease of access to published articles via publically available websites. For this thesis, published articles were acquired via the UCL library services subscription. Extraction of data however remains a manual process following a review of the article into local software identified as being fit for purpose.

3.2 Data management & analysis

Clinical trial data are published in relatively standard format owing to the integrity of the editorial process in ensuring that the study has been conducted to a certain standard. Demographic data relating to characteristics of the population at baseline is generally presented as a summary in table 1 with any specific variations described in the text. Key outcome measures specified *a priori* within the trial protocol are generally presented as a summary in table 2 with appropriate statistical analysis. Important methodological data and additional analyses are frequently published separately as supplementary material which is available as an adjunct to the main article. To ensure integrity in data management, it is important that datasets are carefully reviewed, extracted and analysed. In order to do this, a second independent person is required.

3.3 Evidence Based Medicine

The Cochrane Collaboration is an international organisation whose primary aim is to help people make well-informed decisions about health care by preparing, maintaining and promoting the accessibility of systematic reviews of the evidence that underpins them. The Cochrane Handbook provides guidance to authors for the preparation of Cochrane Intervention reviews. The latest version 5.1.0 (last edited 20 March 2011) is available online at http://handbook.cochrane.org/

The Decision Support Unit (DSU) is a collaboration of Universities with expertise in data synthesis, as commissioned by the National Institute for Health and Clinical Excellence (NICE) to provide research and training resource to support the Technology Appraisal Programme. A series of seven Technical Support Documents (TSD) are published and publically accessible. TSD 7 relates to a reviewers checklist describing points that should be addressed as part of the research plan.(21)

Recommendations from the above were considered as part of the scope, protocol, methodology, assessment of outcomes and translation to clinically relevant findings from the various projects covered within this thesis.

3.4 Systematic Review

Systematic reviews provide one of the highest levels of clinical evidence(22) and is an important part of the hierarchy of evidence (see Figure 2).

Chalmers and Altman defined a systematic review as *"a review that has been prepared using a systematic approach to minimising biases and random errors which is documented in a materials and methods section. A systematic review may, or may not, include a meta-analysis: a statistical analysis of the results from independent studies, which generally aims to produce a single estimate of a treatment effect."(23)*

A systematic review is an evidence-based critical assessment and synthesis of the results of trials or studies. This information can be used to shape medical decision making, to inform policy makers, to keep health care practitioners up to date, and to highlight areas where more research is needed.(24, 25)

It is important that any future research builds on existing theories or knowledge, and that it does not duplicate the work of others. A systematic review provides a comprehensive overview of the extent of research on a particular subject and the level of current knowledge. The findings of a systematic review assist in defining research

objectives to fill in any important gaps as well as determining the strengths and weaknesses of different methodologies available.

3.4.1 Key points

- Reviews are essential tools for health professionals, researchers, consumers and policy makers who want to keep up with the evidence that is accumulating in their field
- Systematic reviews allow for a more objective appraisal of the evidence that traditional narrative reviews and may thus contribute to resolve uncertainty when original research, reviews and editorials disagree
- Meta-analysis (see section 3.5.1) , if appropriate, will enhance the precision of estimates of treatment effects, leading to reduced probability of false negative results, and potentially to a more timely introduction of effective treatments
- Systematic reviews may demonstrate the lack of adequate evidence and thus identify areas where further studies are needed

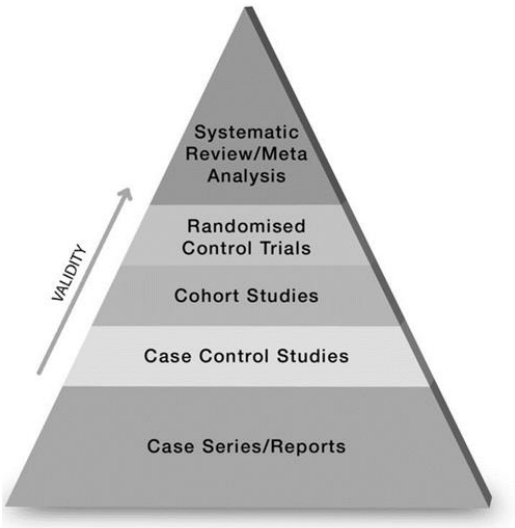


Figure 2: Pyramid of evidence-based medicine (from Murad MH et al.)(26)

3.4.2 Validity

The validity of a systematic review is partially based on a comprehensive literature search. The *Cochrane Handbook for Systematic Reviews of Interventions*, which details the process of preparing Cochrane systematic reviews, defines a comprehensive search strategy as being replicable and thorough and as including a search of various sources.(27)

The key characteristics of a systematic review are:

- A clearly stated set of objectives with pre-defined eligibility criteria for studies
- An explicit, reproducible methodology
- A systematic search that attempts to identify all studies that would meet the eligibility criteria
- An assessment of the validity of the findings of the included studies, for example through the assessment of risk of bias
- A systematic presentation, and synthesis, of the characteristics and findings of the included studies

An important advantage of systematic reviews is that they render the review process transparent. In traditional narrative reviews it is often not clear how the conclusions follow from the data examined. In an adequately presented systematic review it should be possible for the readers to replicate the quantitative components of the argument.

3.4.3 Software

To support the robust preparation and execution of a systematic review, the Cochrane Community make available Review Manager (RevMan), a software that is mandatory for any author of a Cochrane Review. RevMan, currently on version 5.3.5, has been developed through a continuous process of consultation with users and methodologists to ensure that data collection and analytical methods included are of the highest standard. RevMan 5 is free to access with the intention that reviews are for academic purpose.(28)

3.4.4 Search Strategy

A comprehensive search strategy involves the development and implementation of a strategy that will lead to the identification of relevant published and (where required) unpublished studies from indexed biomedical databases and grey literature sources. It is recommended that a comprehensive search strategy include the use of two or more databases (with consideration of the unique contributions of each database); the hand-searching of the bibliographies of selected articles, conference proceedings, and abstracts; and personal communications with researchers.(29, 30) These steps are advocated to minimise selection bias.

3.4.5 Reporting Quality of Randomised Evaluations

To address the suboptimal reporting of meta-analyses in the past, an international group developed a guidance called the QUOROM statement (QUality Of Reporting Of Meta-analyses), which focused on the reporting of meta-analyses of RCTs. Recently, a revision to this guideline has been published, renamed PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses), which has been updated to address several conceptual and practical advances in the science of systematic reviews.(31) Key changes are the inclusion of a protocol from the outset, which has been a standard part of the Cochrane Collaboration review process, and the evolution of terminology to encompass both systematic reviews and meta-analyses. The definitions used within this updated statement have largely been adopted from those used by the Cochrane Collaboration.

Steps in conducting a systematic review

1. Formulate the research question
2. Define inclusion and exclusion criteria
   - 2.1. Participants
   - 2.2. Interventions and comparisons
   - 2.3. Outcomes
   - 2.4. Study design and methodological quality
3. Locate studies (develop search strategies considering the following sources)
   - 3.1. The Cochrane Controlled Trials Register (CCTR)
   - 3.2. Electronic databases and trials registers not covered by CCTR
   - 3.3. Checking the reference lists
   - 3.4. Hand-searching of key journals
   - 3.5. Personal communication with experts in the field
4. Select studies
   - 4.1. Have eligibility checked by more than one observer
   - 4.2. Develop strategy to resolve disagreements
   - 4.3. Keep log of excluded studies, with reasons for exclusions
5. Assess study quality
   - 5.1. Consider assessment by more than one observer
   - 5.2. Use simple checklists rather than quality scales
   - 5.3. Always assess concealment of treatment allocation, blinding, and handling of patient attrition
   - 5.4. Consider blinding of observers to authors, institutions and journals
6. Extract data
   - 6.1. Design and pilot data extraction form
   - 6.2. Consider data extraction by more than one observer
   - 6.3. Consider blinding of observers to authors, institutions and journals
7. Analyse and present results
   - 7.1. Tabulate results from individual studies
   - 7.2. Examine forest plots
   - 7.3. Explore possible sources of heterogeneity
   - 7.4. Consider meta-analysis of all trials or sub-groups of trials
   - 7.5. Perform sensitivity analyses, examine funnel plots
   - 7.6. Make a list of excluded studies available to interested readers
8. Interpret data
   - 8.1. Consider limitations, including publication and related biases
   - 8.2. Consider strength of evidence
   - 8.3. Consider applicability
   - 8.4. Consider numbers-needed-to-treat to benefit / harm
   - 8.5. Consider economic implications
   - 8.6. Consider implications for future research

Figure 3: Steps in conducting a systematic review

3.4.6 Bias

A bias is a *systematic error*, or deviation from the truth, in results or inferences.(27) The introduction of bias can influence the results of systematic reviews and lead to inaccurate pooled estimates of the effect measures. Biases can operate in either direction or at different levels of magnitude; different biases can lead to underestimation or overestimation of the true intervention effect. The greater the rigor in conduct of the systematic reviews, the more likely it will yield results that are closer to the truth. A number of factors can alter the estimate of an intervention's effectiveness following a systematic review, such as:

- Language restriction
- Publication status
- Sources of funding
- Outcome reporting
- Databases searched
- Quality of study

It is important to acknowledge that bias is not synonymous with imprecision. Bias refers to a *systematic error*, meaning that replication of the same process would reach the same conclusion. Imprecision however relates to *random error*, where replication of the same process would result in different outcomes.

3.4.6.1 Language Restrictions

The English language is generally perceived to be the universal language of science.(32, 33) The top 10 international medical journals in 11 medical specialities (measured by impact factor) are English-language publications. The exclusive reliance, however, on data that are published in English, and that are used as the basis of systematic reviews of health care interventions, may not result in an accurate representation of existing evidence. Therefore, excluding publications in a language other than English (LOE) may lead to erroneous conclusions due to the introduction of language bias.

A review by the Canadian Agency for Drug and Technologies in Health (CADTH) could find no evidence of a systematic bias from the use of language restrictions in systematic reviews or meta-analysis of conventional medicines.(34) However, it was suggested that more research is needed, particularly in specific medical specialities, to better understand the role of language restriction. It is possible that the inclusion of publication in LOE may increase the external validity of the report where publications that are written in LOE are known to be influential. As such, it was a recommendation that systematic reviewers should include the search of foreign language studies when resources and time are available.

3.4.6.2 Quality of Study

The quality of study reporting is dependent upon when and where it is published. Recently published studies are subject to greater rigour than those previously with greater requirements imposed on trial protocols as well as conditions and reporting of the study. Advances in the respective field of the study topic also improve the quality of subject inclusion and assessment criteria. Those that are published will have undergone a comprehensive peer review process and therefore brought to a high standard worthy of circulation to the scientific community.

3.4.6.3 Publication Status

Publication bias is a well-known phenomenon whereby positive results have a better chance of being published and in journals with a higher impact factor. Conclusions based on published studies alone therefore have the potential to be misleading, however this should be weighed against the quality of study reporting for those that are unpublished (i.e. the absence of a formal peer review assessment renders the study and its findings to be unvalidated and therefore itself has the potential to be misleading).(35)

It is not uncommon to see industry-funded trials published more routinely than those which are not. Journals that are marketed as open-access charge authors a publication fee which also acts as a barrier for non-industry sponsored publications, particularly if the

outcomes are negative. Almost all journals earn considerable revenue from sales of reprints of articles; this may also be a potential conflict of interest that affects publication patterns.

3.4.6.4 Risk of Bias Assessment

Risk of bias should be assessed using the Cochrane Risk of Bias Tool.(36) The purpose of this is to assign a judgement of high, low or unclear risk of bias for each of the following domains: selection bias (randomisation sequence and allocation concealment), performance bias (blinding of participants and carers), detection bias (blinding of outcome assessment), attrition bias (due to dropouts and exclusions), and reporting bias (selective outcome reporting). The assessment should be completed by one individual and then independently verified by another to ensure robust reporting.

3.4.6.5 Source of Funding

The pharmaceutical industry plays a vital role in financing the research required to develop new medicines. However, corporate financing of clinical research, which may be perceived as providing incentives for investigators and / or control over the study itself, may create conflicts of interest that can bias study results, for example methodological or reporting bias in favour of the investigational treatment yielding positive results that otherwise would not be seen.(37)

3.4.7 Generalisability (External Validity)

The generalisability is concerned with the extent to which the findings of a study can be applied to individuals beyond the sample. Many studies involve samples rather than a whole population. Studies are also focussed on a single location or a small number of areas, but there may be a strong argument that the findings have wider relevance.

The most important issues that determine the generalisability of study findings are that the sampling strategies, procedures and sizes, response rates, and completeness of data.

3.5 Meta-analysis

3.5.1 Concept

Several studies are often published, investigating the efficacy and safety of the same medicine. Although the optimal way in which to determine the effects of a treatment is via a randomised controlled trial, undertaking one including a sufficiently large number of subjects becomes both overly complicated and costly. A frequentist meta-analysis [henceforth referred to as meta-analysis] can be conducted to pool data concerning an endpoint of interest to estimate the overall point estimate and variance [often displayed as a pooled odds ratio or weighted / standardised mean difference with associated confidence intervals]. A meta-analysis is the quantitative analysis of two or more independent studies for the purpose of determining an overall effect and for exploring reasons for variation between study results. Like systematic reviews, meta-analyses are the highest level of evidence based medicine (see Figure 2). It is important to ensure that studies that are pooled together must have sufficiently homogenous characteristics.

The merits of supporting a systematic review with a meta-analysis of the data include:

- To increase power: Power is the chance of detecting a real effect as statistically significant if it exists. Many individual studies are too small to detect small effects, but when several are combined there is a higher chance of detecting an effect.
- To improve precision: The estimation of an intervention effect can be improved when it is based on more information.
- To answer questions not posed by the individual studies: Primary studies often involve a specific type of patient and explicitly defined interventions. A selection of studies in which these characteristics differ can allow investigation into the consistency of effect and, if relevant, allow reasons for differences in effect estimates to be investigated.

- To settle controversies arising from apparently conflicting studies or to generate new hypotheses: Statistical analysis of findings allows the degree of conflict to be formally assessed, and reasons for different results to be explored and quantified.

Meta-analyses are increasingly being used to inform on clinical practice in healthcare as they are able to increase the power of the results observed in small RCTs providing a robust and unbiased account of the literature with more statistical certainty. It is important that a meta-analysis is not performed if there is significant heterogeneity between studies as it has the potential to inflate small study bias and therefore mislead on outcome measures.

3.5.2 Model type

The most frequently used statistical model used for meta-analysis is the Frequentist method. This assumes that unknown parameters are fixed constants, with probability defined by use of limiting relative frequencies.

A meta-analysis may be based on a fixed-effect or random-effects model. In a fixed-effect model, it is assumed that the true effect of a treatment is the same value in each study (i.e. homogenous or *fixed*) across the patient population enrolled, with the differences between each study result being due solely to chance. An example of a fixed-effect statistical measure is the *Mantel-Haenszel method*. A fixed-effect model takes the inverse variance of the estimates as weights, and interpretation relies on an assumption of a common effect underlying every study. For example, if the effect size index used is the *d* value, the fixed-effect model assumes that the population value of *d* is the same in all studies included in the meta-analysis.

In a random-effects model, the above assumption is not made, i.e. the treatment effects for the individual studies are assumed to vary around some overall average treatment effect.(38) A random-effects model incorporates the underlying among-study variation of effects into weights, which allows for the possibility that the population parameter varies from study to study.(39)

The approach used affects the estimated overall effect and its corresponding 95% confidence interval, therefore it is important to understand this when reporting the results. There is currently no consensus on whether it is more appropriate to use the fixed- or random-effects models, hence both are usually performed by the author and an explanation provided as to the choice made when reported the findings. This choice should reflect knowledge of the constituent data included within the analysis.

3.5.3 Challenges

Although meta-analysis is an objective process of synthesising studies, there are subjective decisions to make in the process. The results may be biased if certain aspects are not handled appropriately, for example (1) robust inclusion / exclusion criteria; (2) suitable statistical model and methods; (3) complete reporting with acknowledgement of strengths and limitations. These points should be clearly described to enable the reader to understand the process and assumptions made.

3.5.4 Heterogeneity

An important aspect of both systematic review and meta-analysis is an assessment of the consistency of treatment effect across the primary studies identified against the criteria. Identified trials will have not been undertaken against the same protocol, and as such there may be variations between patients groups, clinical setting, concomitant care / medication, and the method of delivery of the intervention. The possibility of variability, or heterogeneity, between the results of different trials should therefore be examined. See section 3.5.7 for further detail.

3.5.5 Summary statistics

3.5.5.1 Odds ratios

An odds ratio (OR) is a measure of association between an exposure and an outcome.(40) The OR represents the odds that an outcome will occur given a particular exposure, compared with the odds of the outcome occurring in absence of the exposure. The OR is therefore used for a dichotomous endpoint where an event either will or will not occur.

The OR is derived by dividing the number of subjects with an event by the number who did not have the event. The OR can be calculated using the simple principle [A x D / B x C] within Table 1 below. The accuracy of the OR is based on the size of the sample, where in general, the larger the sample size the more robust the estimate is. For this reason, it is also conventional to calculate the 95% confidence intervals for the OR.

Table 1: Calculating an odds ratio

	Response	No Response
Active Treatment	A	B
Placebo	C	D

Equation 1: Calculating the 95% confidence interval for an odds ratio

$$95\%\ \text{CI of ln(OR)} = \ln(OR) \pm 1.96\ (1/A + 1/B + 1/C + 1/D)^{0.5}$$

An OR > 1 proposes that the exposure is associated with higher odds of achieving the outcome, whilst an OR < 1 proposes that the exposure is associated with lower odds of achieving the outcome. For example, if a new medicine is being assessed compared to placebo for its effect on a positive endpoint (i.e. efficacy), the OR would need to be greater than one to demonstrate it being more effective than placebo.

The 95% confidence interval (CI) is used to measure the precision of the OR. A large CI compared with the OR estimate indicates a low level of precision of the OR, whereas a small CI indicates a

higher level of precision. The 95% CI does not report statistical significance, however, the 95% CI is often used as a proxy for the presence of statistical significance if it does not overlap the null value, i.e. the OR and its 95% CI are all greater than 1.

3.5.5.2 Hazard Ratio (HR)

The hazard ratio (HR) is reported to describe an outcome where the parameter of interest is affected by time. Being a ratio however, the value only informs on the extent by which the exposure reduces (or increases) the risk of the outcome from occurring at a given time point, and does not inform on time-dependent effect; in these circumstances a time-to-event curve, such as Kaplain-Meier, would be required.

The interpretation of HR is the same as for OR, keeping in mind that the outcome relates to a time-sensitive endpoint.

3.5.5.3 Mean Difference (MD)

The mean difference (MD), also known as difference in means, is a measure of the absolute difference between the mean values in two difference groups. The use of absolute difference also prevents issues with negative numbers, e.g. the difference between -1 and 1 is 2. MD enables an assessment of continuous data. See Figure 4 for detail.

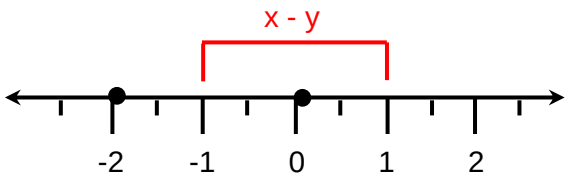


Figure 4: Pictorial representation of mean (absolute) difference

In situations where clinical trials have reported a common endpoint but utilised different units of measurement it is still possible to compare the means between groups using standardised mean difference (SMD). The SMD is a way to measure effect size, it

standardises test results to enable comparison. Depending on the software used the SMD may be calculated via a pooled standard deviation (Cohen's d) or via a weighted and pooled standard deviation (Hedges' g), the latter referred to as Weighted Mean Difference (WMD).

For continuous variables, in order to calculate an MD it is required to have the mean, standard deviation and sample size of each group. For the WMD calculation, the weight given to each study (i.e. how much influence the study has on the overall results) is determined by the precision of its estimate of effect or the number of participants included.

3.5.5.4 Number Needed to Treat (NNT)

The number needed to treat (NNT) is a way of reporting the results for which the outcome measure is binary, i.e. success or failure. In such an analysis, the comparison of one treatment with another provides an NNT which is a value identifying the number of patients needed to be treated with the new intervention to achieve one more success that would have been obtained by treating with the standard intervention.(41) Cook and Sackett (1995) proposed that NNT is clinically easier to interpret than odds ratio, risk reduction, mean difference, etc.(42) The NNT is also more powerful as it provides a measure of the clinical significance of a treatment effect over and above statistical significance.(43)

NNT is calculated as the reciprocal of the difference between the proportion of success on the new treatment and the proportion of success on the old treatment. For example, the proportion of patients with an infection successfully treated with a new antimicrobial is 0.533 and with the existing antimicrobial is 0.133; therefore the NNT is 1 / (0.533-0.133) = 1 / 0.4 = 2.5. This means, for every 2.5 patients treated with the new antimicrobial we will have one more successful outcome. The smaller the NNT, the better the new treatment is in comparison to the existing treatment. In general, the NNT would always be positive; however it is possible to calculate a negative result which would indicate that the new

treatment is more harmful / less successful than the existing treatment. As with OR, 95% CI are also calculated to indicate precision of the estimate.

Where the endpoint under assessment is a negative one, e.g. a safety endpoint, the analysis would be termed number needed to harm (NNH) where the expectation would be for the new treatment to have a larger NNH i.e. treat more patients to result in one more negative outcome.

3.5.6 Displaying the Data

3.5.6.1 Forest Plot

A forest plot is the graphical representation of a meta-analysis. Each study is represented by a black square and a horizontal line which corresponds to the point estimate (median) and the 95% confidence interval, plotted according to the same outcome measure, e.g. odds ratio or mean difference. The area of the black square reflects the weight (contribution to the overall result based on its population size) of the study within the meta-analysis. The 95% confidence intervals would contain the true underlying effect in 95% of occasions if the study were repeated again and again. Each line is accompanied by a reference (author, date, number of subjects in the experimental arm with mean treatment effect, number of subjects in the control arm with mean treatment effect, weight (%), and / or study outcome measure).

Below the studies, an overall result will be estimated indicating the pooled (weighted average) outcome measure, represented as a diamond. The width of the diamond may be reflective of the 95% confidence interval if a horizontal line is not reported. The p-value indicates the level of statistical significance. If the estimate or its 95% CI (or the diamond in the case of the pooled estimate) do not touch the line of unity, the difference found between the two groups will be regarded as statistically significant i.e. $p<0.05$. See Figure 5 for detail.

A logarithmic scale is often used when plotting odds ratio data on a forest plot, primarily because it is easier to represent the graphics in this format. In this format, an effect of 0.5 and 2 represent ratios of the same magnitude but in opposite directions, being equidistant from 1, the null effect. This allows the presentation of data (e.g. confidence intervals) using this scale to be plotted in a more symmetrical format around the point estimate.

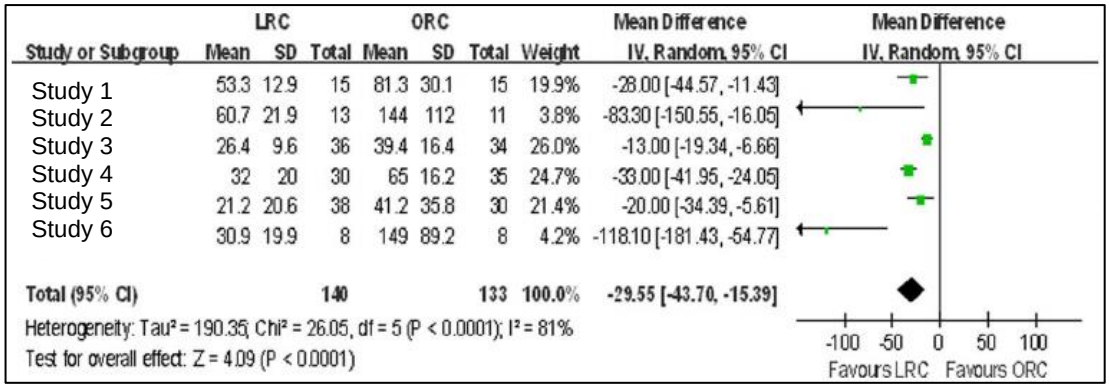


Figure 5: Forest plot (example) of the treatment effect estimated from individual studies.

The overall weighted pooled outcome estimated lies to the left of the line of unity (for an outcome where a reduction indicates a better treatment response) indicating that the interventional treatment (LRC) is more effective than the control treatment (ORC). As the entire diamond lies to the left of the line, the pooled treatment effect is considered statistically significant.

3.5.7 Examination of results

3.5.7.1 Funnel Plot

Funnel plots are used to examine bias in the results of meta-analysis. The plot is a scatter of effect estimates from the individual studies against a measure of each study's size or precision. The standard error of the effect estimate is often used as the measure of study size and plotted on the vertical axis with a reversed scale that places the larger, more powerful studies towards the top. The effect estimates from smaller studies should scatter more widely at the bottom. In the absence of bias and between study heterogeneity, the scatter will be due to sampling variation alone and the plot will resemble a symmetrical inverted funnel (see Figure 6). A triangle centred on a fixed effect estimate and extending 1.96 standard errors either side will include approximately 95% of studies if no bias is present and the assumption is valid. If smaller or non-significant studies are less likely to be published, trials in the bottom left hand corner (when a desirable outcome is being considered) of the plot are often omitted, creating a degree of asymmetry in the funnel (publication bias). Although a subjective tool, the funnel plot is generally considered a good explanatory tool for investigating publication bias.(44)

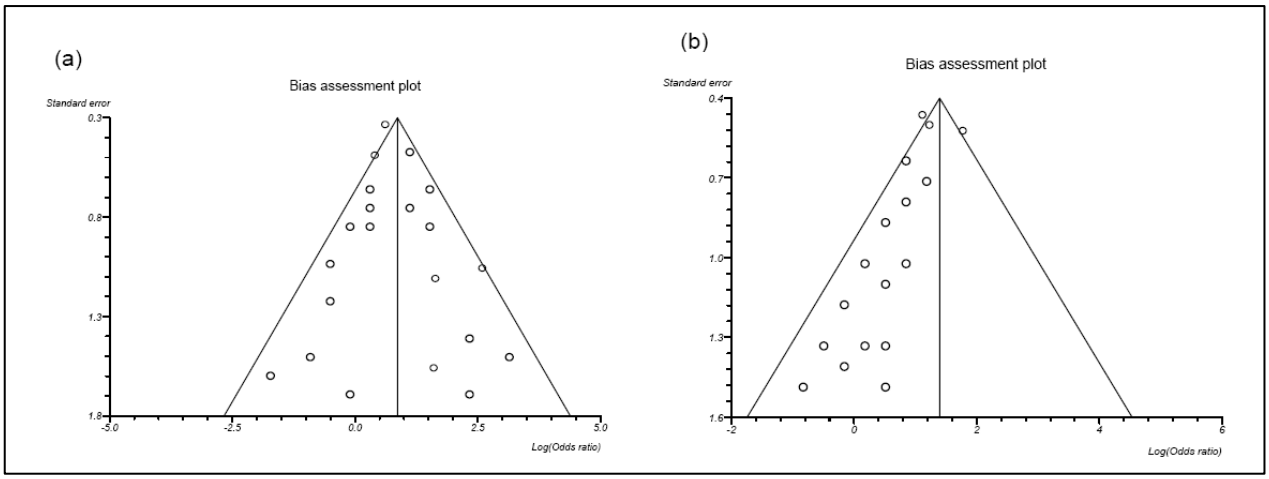


Figure 6: A scatter plot (funnel plot)

Treatment effect estimated from individual studies on the horizontal axis against the standard error (proxy for study size / precision) on the vertical axis. The outer dashed lines indicate the triangular region within which 95% of studies are expected to lie in the absence of bias or heterogeneity. The solid vertical line corresponds to the line of unity and therefore no intervention effect. Image (a) represents a symmetrical plot in the absence of bias whilst (b) depicts an asymmetrical plot in the presence of bias. Image reproduced from Sterne JAC et al. 2011.(44)

3.5.7.2 L'Abbé Plot

L'Abbé plots are used to illustrate treatment effect for one intervention compared with another.(45) The l'Abbé plot visually expresses within group variations in observed results through a plot of the event rate in the treatment group on the vertical axis and the control group on the horizontal axis.(45) Trials in which the experimental treatment is better than the control will be in the upper left quadrant of the plot, between the y-axis and the line of equality. If the experimental treatment is no better than the control, then point estimate will fall on the line of equality, and if the control is better than experimental then the point will be in the lower right of the plot, between the x-axis and the line of equality. The dashed line indicates the estimated effect. The size of the points is drawn proportional to the study size. The l'Abbé plot helps visualize evidence for small study bias of which publication bias is a potential cause.(46, 47) See Figure 7 for detail.

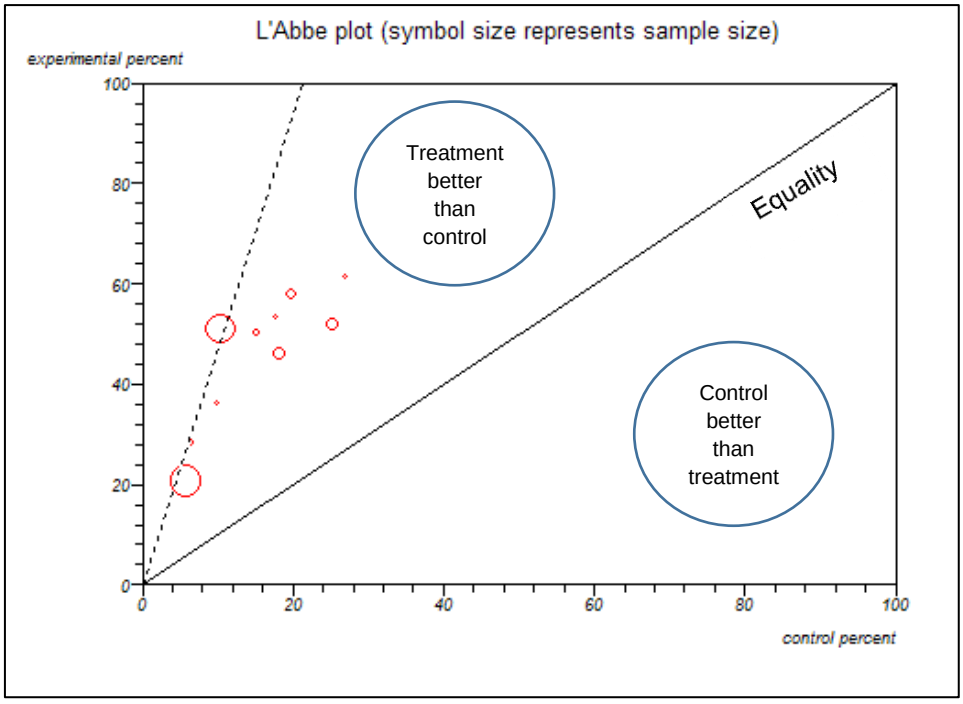


Figure 7: L'Abbé plot

Plot indicates the event rate in the experimental (intervention) group against the event rate in the control group, as an aid to exploring heterogeneity of effect estimate within a meta-analysis.

3.5.7.3 Cochran Q statistic

The Cochran Q is calculated as the weighted sum of squared differences between individual study effects and the pooled effect across studies, with the weights being those used in the pooling method which forms part of the DerSimonian-Laird random-effects model.(48)

This statistic is based on the squared difference between the estimated treatment effect in one trial (T_i) and the overall estimated treatment effect (T) weighted by the inverse of the estimated variance of the treatment effect in T_i (W_i); where k is the number of studies being combined.

Equation 2: Calculation of the overall estimated treatment effect

$$T = \frac{\sum W_i T_i}{\sum W_i}$$

Equation 3: Calculation of the Cochran Q statistic

$$Q = \sum_{i=1}^{k} W_i (T_i - T)$$

Under the null hypothesis of homogeneity, this statistic follows a chi-squared distribution with k-1 degrees of freedom. When the value of the Q-statistic is large (that is, overall the trial effects are far from the mean effect, taking into account sampling variance) this is indicative of substantial heterogeneity and the null hypothesis of homogeneity is rejected. If there is no evidence of heterogeneity, the value of Q should be approximately equal to k-1.

3.5.7.4 Higgins I-squared statistic

The I-squared statistic is a more intuitive, simpler expression compared with Cochran Q, which describes the proportion of variation across studies that is due to inconsistency rather than chance.(48) The Cochran Q statistic must be calculated first. As the statistic is expressed as a percentage, any negative values should be read as equal to zero. Values of I-squared equal to 25%, 50%, 75% and 100% represent low, moderate, high, and very high degrees of heterogeneity, respectively.

Equation 4: Calculation of the I-squared statistic

$$I^2 = 100\% \times \frac{Q\ (k-1)}{Q}$$

The magnitude of heterogeneity in a random-effects model can be quantified by calculating a point estimate of the among-study variance of true effects; tau-squared. If there is no variance between studies, the tau-squared value would be expected to be low (or zero). Where a value greater than 1 is reported, this would suggest the presence of substantial statistical heterogeneity.

3.6 Network Meta-analysis (NMA) and Mixed Treatment Comparison (MTC)

3.6.1 Concept

Like the [frequentist] meta-analysis, a network meta-analysis (NMA) is where the results of several quantitative studies are pooled together if homogenous characteristics are selected. However, a NMA offers a set of methods to visualise and interpret the wider picture of the evidence and to understand the relative merits and / or disadvantages of multiple interventions that may not have been compared directly, provided a 'loop' of direct comparisons can be made somewhere in the generated 'network'.(49) NMA is also commonly referred to as multiple or mixed treatment comparison. NMA is commonly based on Bayesian methods, considered an alternative approach to the Frequentist method. Under Bayesian methods, the parameters are treated as random variables with probability defined as 'degrees of belief' i.e. the probability of an event is the degree to which you believe that the event is true.

The results an NMA estimates for each pairwise comparison is achieved by combining all of the 'direct evidence' (evidence based on head-to-head comparisons between interventions made within individual studies) with all of the 'indirect evidence' (comparisons between interventions inferred from the network via common comparator interventions).(49, 50) In effect, indirect evidence comparing the effect of interventions A and B can be inferred from the direct evidence provided by a trial comparing A with C and a trial comparing B with C. NMA thus enables estimation of relative intervention effect estimates for every pair of interventions, regardless of whether or not they have been compared directly in a RCT. It also enables the ranking of treatments according to the probability that each is the best, or worst, for a given outcome.

3.6.2 Introduction

When multiple interventions have been used and compared for the same disease and outcomes, network meta-analysis (also commonly referred to as multiple treatment comparison meta-analysis or mixed

treatment meta-analysis) offers a set of methods to visualise and interpret the wider picture of the evidence available and to better understand the relative merits and / or harms of the multiple interventions available.(49)

The use of network meta-analysis has advantages over the traditional frequentist meta-analysis, as the methodology utilises the strengths from indirect evidence to increase certainty about the effect of all treatment options and also allows for estimation of comparative effects that have not been investigated in a head-to-head manner in randomised controlled trials.(50)

Network meta-analysis has quickly gained popularity amongst clinicians, guideline developers, and health technology assessment agencies as part of the review of new treatments in the context of previously available evidence and treatments.(51) For many comparisons, the network meta-analysis may yield more reliable and definitive results than would a pairwise meta-analysis.

Despite the increasing popularity and widespread use of network meta-analysis, an element of resistance from clinical practitioners remains, likely attributed to poor understanding of the certain methodological and interpretational aspects of this technique. Specific areas that require further elucidation to ensure high quality synthesis of evidence in the setting of multiple treatment options include: the strength of evidence included; risk of bias for each trial and treatment comparison included within the treatment network; tools and assessments in detecting and / or exploring heterogeneity within and between treatment comparisons that have not been conducted in a head-to-head manner; interpretation of the statistical models used and the effect measures estimated by the model.(52) The authors of this US Hospital Evidence-based Practice Centre identified 42 network meta-analyses used as part of evidence synthesis by government agencies including health technology assessment panels, of which the majority used Bayesian methods (80.9 percent). The Bayesian analyses used either non-informative priors or did not report detail about priors used. Surprisingly, data evaluation regarding

convergence, heterogeneity, and inconsistency were not consistently reported.

3.6.3 Methodology: Bayesian Statistical Analysis

Markov Chain Monte Carlo (MCMC) simulation (see section 3.6.6.2), using Gibbs sampling, is an algorithm used to generate a sequence of samples from a joint probability distribution of two or more random variables and is particularly well adapted to sampling the treatment effects (posterior distribution) of a Bayesian network. This approach was used to generate the posterior distribution for each OR of interest. The median of the posterior distribution was taken as the point estimate and the 2.5th and 97.5th centiles provided the 95% credible interval (CrI).

In general, Bayesian CrI and frequentist CI are non-interchangeable as CrI incorporate problem-specific contextual information from the prior distribution whereas CI are based solely on the on the data.(53) The interpretation of Bayesian CrI is that the posterior probability that the parameter lies within the CrI is 95%. For analyses where there are no beliefs on the baseline data or the effect size in advance of the analysis a non-informative prior is entered (i.e. a mean of zero and variance of 10,000). This also permits a more direct comparison with a frequentist meta-analysis which does not have a capability to include a prior. Twenty thousand iterations are traditionally used for each chain in the Bayesian analysis following a burn-in of 20,000 in order to reduce the occurrence of heterogeneity, variation or spurious findings.

3.6.4 Application of an MTC

MTC is an extension of pairwise meta-analysis that allows a comparison of treatments that have been compared in a connected network of treatment comparisons, even if all the treatments have not been directly compared in a head-to-head manner in randomised clinical trials.(28, 196, 197) The strength of the MTC approach is that the estimation of the relative effect between two treatments uses all the information available from the network of evidence including direct

comparisons and indirect comparisons without breaking randomisation.

MTC analyses are an important tool for health economists, statisticians, and decision modellers interested in the extension of pair-wise meta-analysis to network meta-analysis (i.e. indirect treatment comparisons and mixed treatment comparisons). It is also an essential tool to gain an in-depth understanding of statistical modelling as part of evidence synthesis in the context of clinical effectiveness or economic evaluation.

Bayesian methods may be used to statistically combine evidence from networks of trials whilst integrating statistical estimation within a probabilistic modelling framework. Assumptions made in the underlying pair-wise comparisons used for meta-analysis are critically appraised and examined as well as identification and management of potential heterogeneity and inconsistency.

3.6.5 Network Geometry

When conducting and interpreting a network meta-analysis it is important to understand the evaluation of its geometry. That is, which of the treatments (referred to as nodes within the network) have been compared to another in a head-to-head manner in randomised controlled trials, which of the treatments are connected via a statistically generated parameter (connected indirectly through one or more ‘common comparators’), and what is the level of evidence informing each comparison. By examining the connections between each of the interventions included in a graphical manner, the reader of the analysis can determine how strong the evidence is for the treatment as a whole and for the individual comparisons, whether specific comparisons are over-represented or under-represented, and whether the network is well connected. The better connected a network is, the more reliable the estimates it provides will be.

The strength of a network and therefore the analysis it generates is dependent upon its geometry (degree of diversity, co-occurrence, homophily, and auto-looping) and appropriate assumptions to avoid

statistical and conceptual heterogeneity. Variations of network constructs are illustrated in Figure 8.

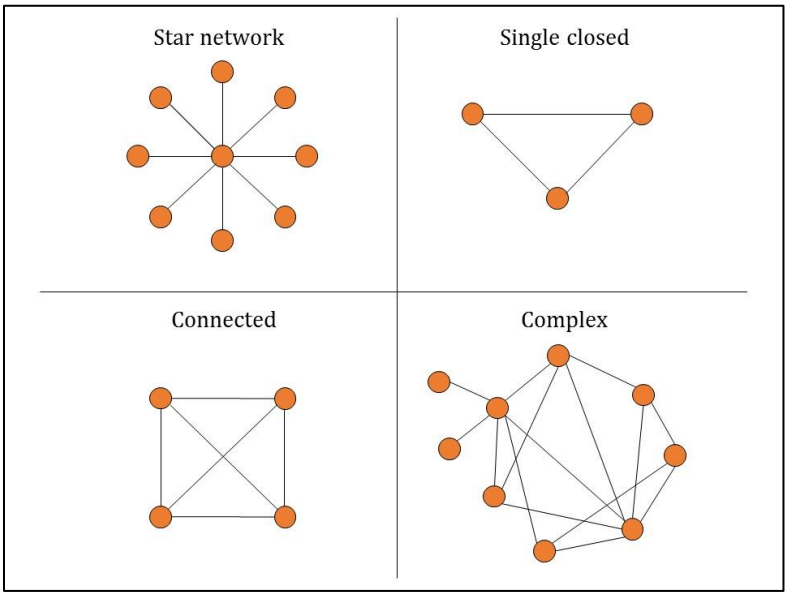


Figure 8: Examples of network geometry that form a network meta-analysis

Another method of determining how strong the evidence is for the treatment network is to provide a comparison of the estimates generated for two treatments compared in a head-to-head manner from a frequentist meta-analysis to the estimate generated by the network meta-analysis. Figure 9 provides an illustration of the 'complex' geometry of a treatment network, which includes 34 randomised pairwise comparisons of which warfarin (n=20), aspirin (n=16) and placebo (n=12) have the most links. Overall, 45 possible pairwise comparisons can be made between the nine treatments. Of these, 16 comparisons are informed directly by head-to-head evidence, but six of the direct connections have only one trial supporting the evidence. Therefore, a number of the comparisons that have not been directly studied are informed by indirect evidence from only two trials. Data generated by the model for nodes within the network that are not connected should be interpreted with caution.

The diversity and strength of a network are determined by the number of different interventions included, the number of comparison of these interventions that are available, how represented they are within the network, and how much evidence they carry. For networks where there is a severe imbalance in terms of the amount of evidence or published data available for each intervention, there is potential that this will adversely affect power and reliability of the overall analysis.(54, 55) This is due to statistical inferences being made largely from the data provided by one or a few treatments within the network.

Although the key strategy of meta-analysis is to strengthen the estimation of results through the pooling of similar data, evidence procured from small trials may be susceptible to a higher degree of bias (for example, more prominent publication and selective reporting bias) as well as showing spuriously larger treatment effects.

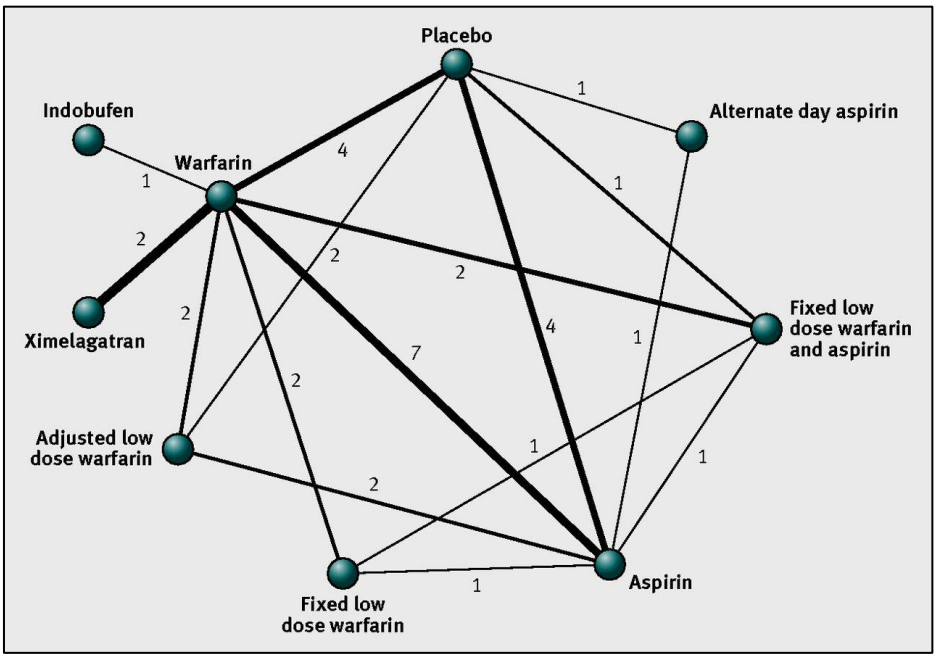


Figure 9: Geometry of a well-connected network of randomised controlled trials

Example illustrates an evaluation of stroke prevention among populations with atrial fibrillation (AF).(56) Circles represent the drug as a node in the network; lines represent direct comparisons using RCTs; the thickness of lines represents the number of RCTs included in each comparison, also represented by the numbers.

3.6.6 Bayesian Statistics

The use of statistical analyses in the 20th century favoured the frequentist statistic; however, a number of flaws in its design and interpretation were being apparent. These largely centred around the estimates being based against a fixed statistic (i.e. the p-value and the 95% confidence intervals). The use of Bayesian statistics has recently seen resurgence as a solution to a number of important problems in medicine being a method to determine the likelihood of the occurrence of an event.

Bayesian statistics is a mathematical procedure that applies probabilities to statistical problems. An important difference compared with Frequentist analysis being the inclusion of prior beliefs (see Figure 10) although this is not necessary. A Prior is acquired from knowledge of the area of study but is independent to the current study, such as knowing what effect a placebo has on an outcome influences the baseline from which an intervention becomes important. An example in lay terms is provided below:

Scenario	Frequentist reasoning	Bayesian reasoning
Situation I have misplaced my phone somewhere in the home. I use the phone locator button on the base of the phone holder to activate the ring on the phone. I am then able to locate the phone by responding to the ring. **Problem** Which area of the home should I search?	**Solution** I hear the phone ringing. I use my ears to identify where the sound is coming from. Therefore, upon hearing the ring, I infer the area of the home I should search to locate the phone.	**Solution** I hear the phone ringing. In addition to using my ears, which will identify where the sound is coming from, I use my *prior* knowledge of where I have previously left or misplaced my phone. I combine my prior knowledge with the ringing to infer the area of the home to search in order to locate the phone.

Another key difference is that the estimates generated report results with a 95% credible interval (CrI) rather than a confidence interval. As this is based on probability, the CrI reports the 95% most credible values, and indirectly guarantees that 95% of the values will like within the range specified, unlike the 95% confidence interval.

There are three key areas where the use of Bayesian methods is most informative (see below). As such, the use of Bayesian methods has been slowly adopted by NICE within their HTA process since 2000.(57)

1. Estimation of clinical effectiveness
2. Modelling of cost-effectiveness
3. Dealing with difficult / complex situations

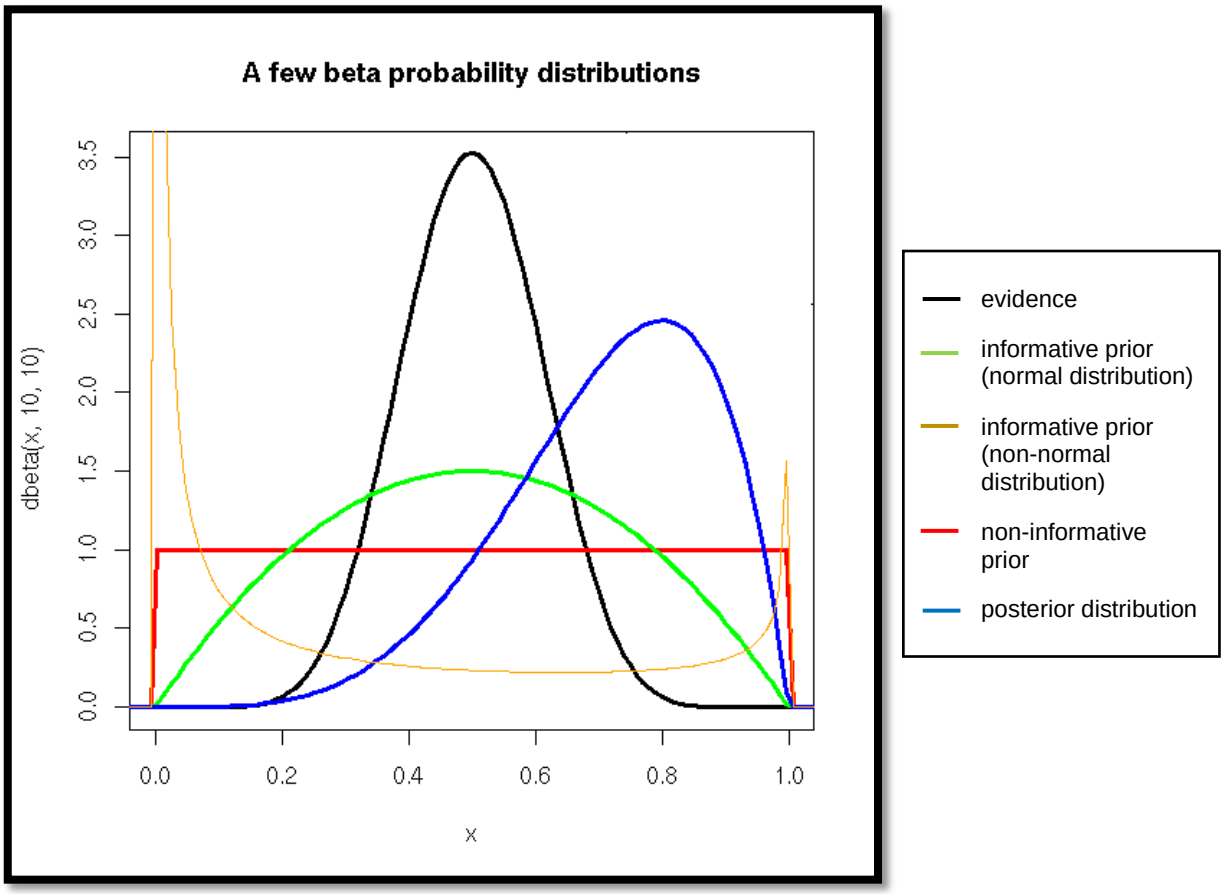


Figure 10: Illustration of the Bayesian distribution (analytical model)

3.6.6.1 WinBUGS

WinBUGS (a Microsoft operating system version of BUGS: Bayesian Analysis Using Gibbs Sampling) is a versatile package that has been designed to carry out MCMC computations for a wide variety of Bayesian models. The software is distributed electronically and available free-of-charge directly from the BUGS project website.(58) This form of logistic regression analysis combines both direct and indirect data without breaking randomisation within a *MCMC* framework. Within this approach, indirect estimates can be combined in large samples if there is no interaction between the treatment effects and the populations or major subgroups in a trial.(59)

WinBUGS can be used in statistical problems as simple as estimating means and variances or as complicated as fitting multilevel models, measurement error models, and missing data models.

WinBUGS fits fixed-effect and multi-level (random-effects) models using Bayesian statistics, whilst Stata fits fixed-effects and limited multi-level models using maximum likelihood or generalised least squares. If no prior or a minimally-informative prior is used then the inferences from the two software should be numerically similar and the 95% confidence interval would be very similar to the 95% credible interval.

The use of WinBUGS software has numerous advantages but also requires care in its use. The ‘burn in’ should be conservatively large to reduce inconsistency in the model outputs; a suggested minimum is 10,000 although 20,000 is conventionally used. Burn in refers to the practice of discarding an initial portion of a Markov Chain so that the influence of the initial value on the posterior inference is minimised. This is done on the assumption that after *t* iterations the chain has reached its target distribution and the earlier portion is discarded, using only the good samples for posterior inference. The value of *t* is the burn in number. Particular care must be taken in checking convergence, with sampling from the posterior distribution, as detailed within the manual. The Monte Carlo error (MCe), which

reflects both the number of simulations and degree of autocorrelation, should be no more than 5% of the posterior standard deviation of the parameters of interest (see section 3.6.8.1).

A step-by-step user guide developed to ensure consistency in application of WinBUGs throughout the thesis is described in section 10.2 (Appendix II: WinBUGS user guide).

3.6.6.2 Markov Chain Monte Carlo (MCMC)

The MCMC method is a general simulation method for sampling from posterior distributions and computing posterior quantities of interest. The MCMC procedure is used to generate a value for each parameter, by sampling randomly from its conditional distribution. This then acts as the 'known' value for that parameter. This process is carried out iteratively. A new parameter value is sampled from the distribution of each parameter in turn, and is used to update the 'known' values for the conditional distribution of the next parameter. The phrase 'Markov Chain' refers to the fact that the procedure is based only on the last sampled values of each parameter, while 'Monte Carlo' refers to the random sampling of the parameter values.

Where used correctly, i.e. the chain is run for a very long time (burn in period of 10,000 iterations), the Markov Chain is guaranteed to converge to the target distribution. The initial values are discarded and the next say 10,000 iterations are recorded corresponding to the posterior distribution. The MCMC method has been successful in Bayesian MTC analysis.

3.6.6.3 Coding

The NICE Decision Support Unit publishes a series of Technical Support Documents on Evidence Synthesis. These describe the analytical approach for meta-analyses in order for the estimates generated to be appropriate for cost-effectiveness analysis and comparative effectiveness research.(60) Within the summary document, guidance on 'good practice' is provided recommending

how the evidence should be presented, including use of network diagrams and content of the results tables. Document 2 specifically outlines the framework for the synthesis of data when taking a Bayesian approach, with WinBUGS program code for a Bayesian analysis using MCMC simulation templates provided.

For each of the analyses performed within this thesis, the WinBUGS program code developed for National Institute for Health and Clinical Excellence (NICE) was adapted for the MTC analysis. (61-63) This coding can synthesise data from pair-wise meta-analysis, multi-arm trials, indirect comparisons and network-meta-analysis. The WinBUGS package computes and presents the analyses in both numerical and graphical outputs.

3.6.6.4 Trade-Off Analysis

An analysis of the strengths and weaknesses of a medicine is traditionally done following a review of its efficacy, safety, convenience and cost relative to alternatives available. This process is conventionally undertaken in a qualitative manner by a visual assessment of a summary table of one agent with another.

In a Bayesian framework, probabilities regarding the distribution of parameters are estimated. In each MCMC cycle, each treatment *'j'* is ranked according to the estimated effect size. The proportion of the cycle in which a given treatment ranks first out of the total gives the probability '*P(j=1)'* that treatment j ranks first i.e. that this treatment is the best among the available treatment options. Similar probabilities are estimated for being the second best, the third best, etc. These probabilities sum to one for each treatment and rank.(64)

The 'Rank Monitor Tool' is a built in function in WinBUGS which is used to estimate the median probability and its 95% confidence interval. The output is presented in table format, therefore in order to simplify its interpretation the values can be presented in graphical format in the form of a rank-o-gram (see Section 3.7.2).

Using the principle that clinicians and patients consider a range of different endpoints when considering if a treatment is the best, the

rank-o-gram is able to accommodate multiple endpoints (each of the efficacy and safety / tolerability analyses performed) that WinBUGs has estimated in order to visualise an overall effect. For the purpose of this thesis, each endpoint included within this analysis is weighted equally and subject to careful interpretation in its reporting.

3.6.7 Displaying the Data

3.6.7.1 Box Plot

A useful diagram to demonstrate values in this manner is a box (and whiskers) plot, as shown in Figure 11. The box is drawn from the lower quartile to the upper quartile; its length therefore provides the interquartile range. The horizontal line in the middle of the box represents the median. The 'whiskers' in the plot mark the extent of the data with 95% confidence. They are drawn on either end of the box to the minimum and maximum values.

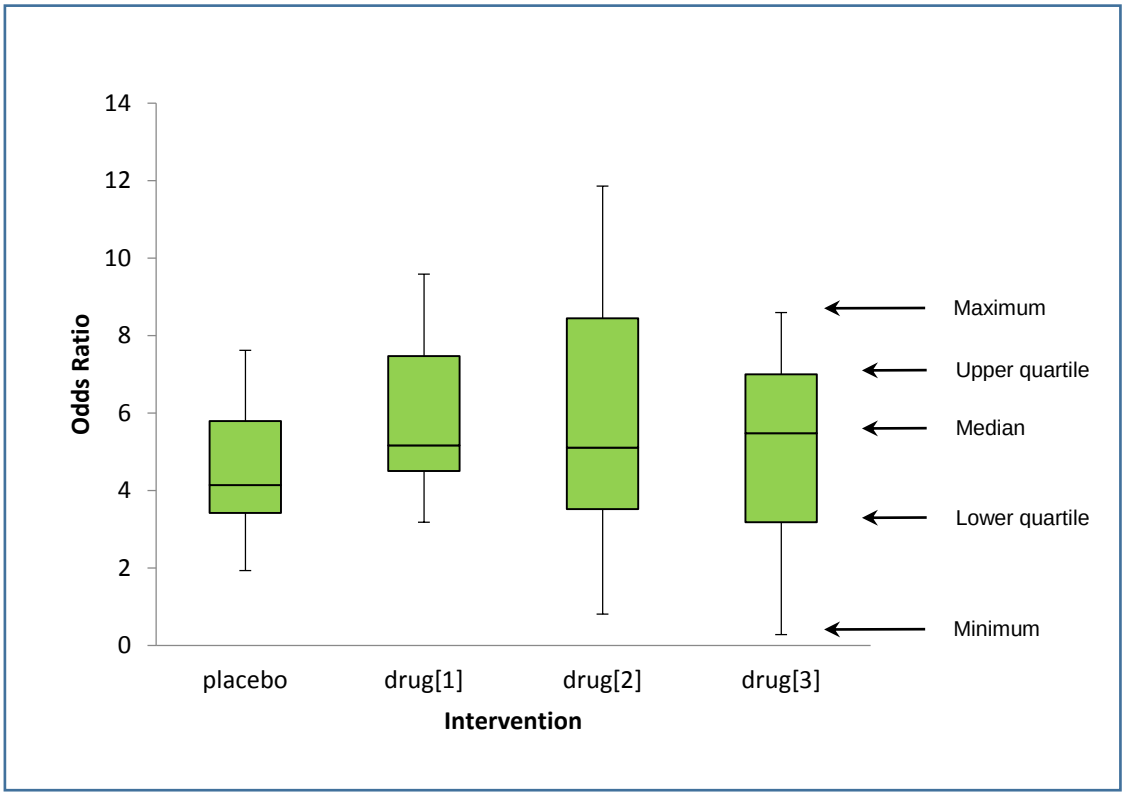


Figure 11: Box and whiskers plot

3.6.8 Examination of NMA Results

3.6.8.1 Convergence

An assessment on convergence should be undertaken to determine if the inferences from the MCMC simulation approach are reliable. Inferences based on non-converged Markov chains can be both inaccurate and misleading. This may be done via visual assessment of trace plots of the posterior distribution. Convergence of Markov chains is deemed to be achieved if plots of the Gelman–Rubin statistics indicate that the widths of pooled runs and individual runs stabilise around the same value.(65) It is difficult to determine conclusively that a chain (simulation) has converged, therefore the visual assessment aims to establish where it has not occurred (see Figure 12). Successful convergence is displayed in the plots as minimal fluctuations (e.g. ± 10) around an alpha point. The visual assessment should be corroborated through assessment of the ratio of the Monte Carlo error (MCe) with its standard deviation (SD) which confirmation of convergence validated with the value being sufficiently small (i.e. MCe/SD <0.05).

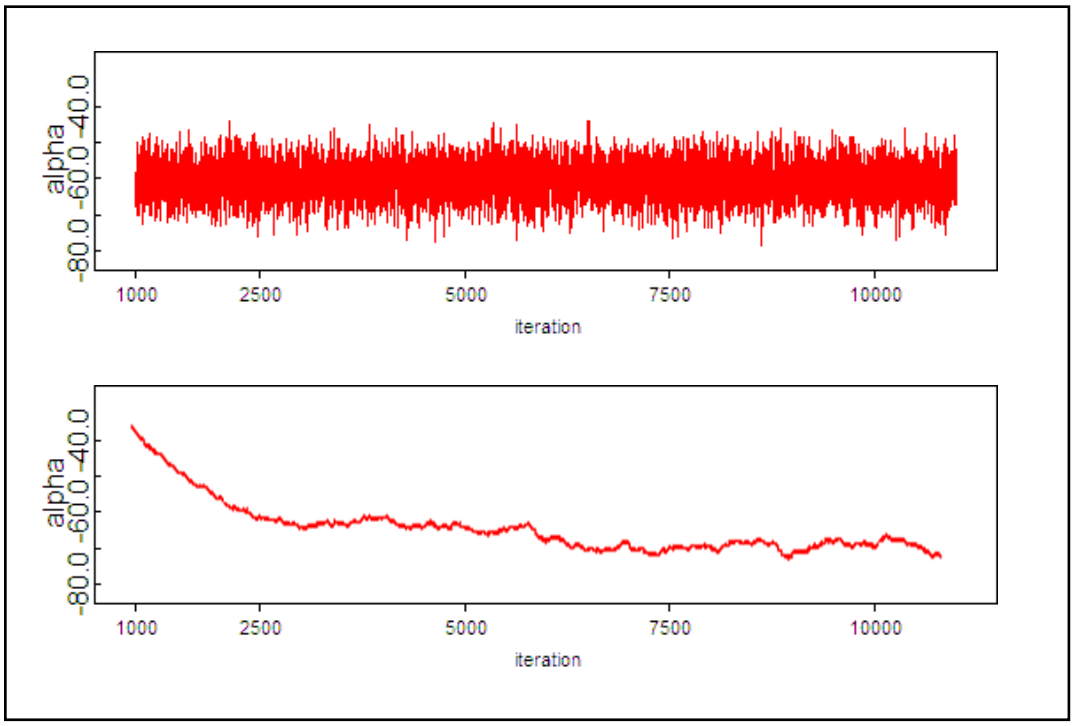


Figure 12: Trace plot (WinBUGs) examples of (top) chain where the convergence looks reasonable, and (bottom) chain which has clearly not reached convergence

3.6.8.2 Inconsistency (Bucher's Test)

The comparison of treatments from different RCTs gives rise to incomplete block structures referred to as a network or mixed treatment comparisons (MTC). In the analysis of outputs from such data structures it is important to determine whether the disparate evidence sources provide consistent information or if the pairwise comparisons are incoherent. This analysis considers whether or not there is any inconsistency between the loops of evidence created of indirect comparisons in the estimation of pairwise comparisons rather than between the individual trials included.(59, 66) The test itself involves an assumption that indirect evidence is consistent with direct comparisons.

The variance is determined, for example, by comparing the difference between treatment A and B in RCT-1 (d_{AB}) with the difference between treatment A and C in RCT-2 (d_{AC}), where treatment A is the common comparator to establish the indirect comparison of treatment B and C. Where there are multiple trials that compared treatment A and B or A and C, the results used will be the estimates generated by the meta-analysis.

Equation 5: Calculation of indirect comparison

$$d_{ind\,BC} = d_{AB} - d_{AC}$$

Equation 6: Calculation of variance of calculated indirect comparison

$$Var\,(d_{ind\,BC}) = Var\,(d_{AB}) + Var\,(d_{AC})$$

The inconsistency test is then determined through an assessment of the calculated difference compared with the actual difference. Based on the assumption that there is no difference between indirect and direct values, the test results should equal zero. The Z-statistic and a corresponding p-value are used to determine if the difference is statistically significant. A Z-statistic value greater than 1.96 (and p-value less than 0.05) are indicative of a statistically significant difference and poor model fit.(67)

3.6.8.3 Goodness of Fit

Like the Cochran Q statistic for Frequentist meta-analysis, the sum of squared deviance (SumDev) residuals depicts how 'good' the data estimated fits within the model through a summary of the discrepancy between observed values and the values expected under the model. An assessment of the 'goodness of fit' of the model should be conducted through analysis of the posterior mean of the sum of the residual deviance contributions of each data point.(62, 68) A good fit would give a result that approximates the number of data points included within the analysis (N); a result which could only occur when the posterior distribution is approximately multivariate normal or the total population included within the analysis (aggregate of all studies) follows a normal distribution. Using the chi-squared distribution, the right-tailed probability can be calculated by a comparison of the SumDev and N, where a p-value less than 0.05 indicates poor model fit.

3.6.8.4 Deviance Information Criterion (DIC)

DIC is equal to the sum of the posterior mean of the residual deviance and the effective number of parameters.(68) The DIC penalises the posterior mean residual deviance (a measure of model fit) by the effective number of parameters in the model (as measure of complexity) and can therefore be viewed as a trade-off between the 'goodness of fit' and 'complexity' of the model. Models with smaller DIC are better supported by the data.

3.6.8.5 Bland-Altman Plot

A Bland-Altman plot (agreement plot) is a method of displaying data to analyse the level of agreement between paired data, or conversely identify systematic difference between measurements or possible outliers. Values plotted should sit around the mean difference and within the boundaries of 95% confidence. The Bland-Altman plot will only graphically demonstrate if a data fall within the intervals of agreement, it is unable to dictate if these limits are acceptable. The limits should be based on clinical relevance.(69, 70)

3.7 Reporting of Trade-off Analysis

3.7.1 NNT / NNH Plot

For analyses where a single efficacy and a single safety / tolerability endpoint are defined as the parameters of interest for a trade-off assessment it is possible to convert these into their corresponding NNT and NNH. Like with a forest plot of odds ratio, a plot of NNT (vertical axis) *versus* NNH (horizontal axis) would be presented on a logarithmic scale. Treatments that feature in the lower right of the plot (i.e. low NNT and high NNH) are considered highly efficacious and well tolerated, ranking highest from the range of treatments investigated. Treatments within this cluster would be identified as the priority for future head-to-head RCT comparative studies.

3.7.2 Rank-o-gram

For analyses where multiple endpoints are defined as the parameters of interest for a trade-off assessment, the best graphical tool to present the probability of each treatment for multiple endpoints in a rank format is a line-graph known as a rank-o-gram. The peak probability (vertical axis) will be associated with a corresponding rank number (horizontal axis).(71) A peak probability at a rank of 1 would usually indicate the best treatment for a positive endpoint (i.e. efficacy), whereas this is likely to indicate the worst treatment for a negative endpoint (i.e. safety). These ‘rank probabilities’ can be estimated directly from WinBUGS. A similar method is available in other software whereby the surface under the cumulative ranking curve (SUCRA) is presented.

3.7.3 Cost-Effectiveness Acceptability Curve (CEAC)

For analyses that have included cost within a formal cost-effectiveness assessment, it is conventional to perform sensitivity analyses to model potential variables in the population characteristics (e.g. age) or model inputs (e.g. cost of treatment). A recognised output of this analysis is a cost-effectiveness plane or acceptability curve, where the probability of an intervention being cost-effective is plotted against a specified value. Treatments that feature on the top left of the plot (i.e. with the highest score of probability at the lowest

willingness to pay) are the most dominant. The critical section of the plot is treatment(s) that appear with the highest probability in the £20,000 to £30,000 range.

3.8 Pharmacoeconomics

Pharmacoeconomics identifies, measures and compares the costs and consequences of drug therapy to healthcare systems and society. A number of economic techniques and evaluations may be used to compare two or more treatments available with the outcome measured in terms of cost and quality. The primary purpose of such economic evaluations is to enable budget-holding decision makers to consider the relative and absolute values of the available resources.

Collaboration and expert support were sought for these aspects of the research projects in order to enhance the frequentist and Bayesian analyses. The additional parameter of cost enabled progress of the findings to guideline development. Further description is provided within each chapter.

3.8.1 Cost-Effectiveness Analysis (CEA)

A cost-effectiveness analysis (CEA) is a form of economic analysis in which consequences of different interventions are measured using a single outcome, usually in ‘natural’ units (for example, life-years gained, deaths avoided, heart attacks avoided or cases detected). Alternative interventions are then compared in terms of cost per unit of effectiveness.(72) Costs included within the analysis may be direct (medicines, staff, equipment, transport), productivity (carers, loss of income), or intangible (pain, suffering, adverse effects).

It is an important component of analysis as it directly relates the financial and scientific implications of providing different interventions to a defined population as identified within the model created. The CEA calculation involves dividing the cost of an intervention in monetary units by the expected health gain measured in natural units (such as lives saved), producing a cost-effectiveness ratio.

3.8.2 Cost-Utility Analysis (CUA)

A cost-utility analysis (CUA) is a form of financial analysis used to guide commissioning decisions, largely utilised by policymakers, with the specific output described in utilities, most commonly quality of life. The most commonly seen and well-known applications of this type of analysis is within the field of pharmacoeconomics, particularly used by NICE within the program of Health Technology Assessments (HTA).

The CUA was developed to address the problem of CEA, which did not allow decision makers to compare the value of interventions for different health problems. As healthcare costs continue to rise the CUA becomes more applicable as it captures the value of improvement in morbidity and mortality.

3.8.3 Cost-Minimisation Analysis (CMA)

A cost-minimisation analysis (CMA) is a method of calculating drug costs to project the least costly medicine or therapeutic modality. This method is used when evaluating the costs of a specific drug compared with another, rather than a variety of alternatives. This method can also only be used when two treatments have been shown to be equivalent in their therapeutic effect (i.e. there is no incremental benefit) and therefore the cost is the only variable.

3.8.4 Quality Adjusted Life Year (QALY)

A quality adjusted life year (QALY) is a measure of health as combination of the duration of life and the health-related quality that an individual experiences within that duration.

3.8.5 Incremental Cost-Effectiveness ratio (ICER)

The cost per QALY, otherwise referred to as incremental cost-effectiveness ratio (ICER) is the ratio between the difference in costs and the difference in benefits of two interventions. The ICER may be stated as (C1 – C0) / (E1 – E0) in a simple example where C0 and E0 represent the cost and gain, respectively, from taking no health intervention action. C1 and E1 would represent the cost and gain,

respectively of taking a specific action. So, an example in which the costs and gains, respectively, are £140,000 and 3.5 QALYs, would yield a value of £40,000 per QALY. These values are useful for policy makers to determine relative priorities when determining treatments for disease conditions. It is important to note that CUA measures relative patient or general population utility of a treatment.

In the UK, NICE is believed to have a threshold of about £30,000 per QALY – roughly twice the mean income after tax. Thus, any health intervention which has an incremental cost of more than £30,000 per additional QALY gained is likely to be rejected and any intervention which has an incremental cost of less than or equal to £30,000 per extra QALY gained is likely to be accepted as cost-effective. This implies a value of a full life of about £2.4 million.

[BLANK PAGE]

4.0 Chapter 4: Research Project One – Antiepileptic drugs for Refractory Epilepsy

4.1 Introduction

This first research project explores the use of the quantitative trade-off methodology in pilot form following Bayesian network meta-analysis. This form of analysis is under-utilised within the evidence-based healthcare setting when attempting to determine the relative benefits and weaknesses of similar interventions. The Bayesian methodology will be undertaken to conduct a network meta-analysis of all published studies comparing anti-epileptic drugs for refractory epilepsy that meet strict criteria from randomised-controlled trials, including of head-to-head design where available, to obtain estimates on their relative efficacy and safety. The use of a Bayesian meta-analysis rather than the traditional frequentist analysis will be suitable for this project as only a limited number of published trials have investigated these agents in a direct head-to-head comparison. The trade-off parameters selected for this analysis include 50% responder rate (reduction in seizure frequency from baseline) vs. withdrawal rate for the efficacy and tolerability endpoints, respectively. Both parameters are weighted equally.

The goal for this project is to understand the strengths and limitations that currently exist with trial design within this therapeutic area, the potential for using this novel analysis, and to influence prescribing at the local specialist centres.

4.2 Literature Review

4.2.1 Epilepsy

Epilepsy is a chronic, complex, neurological condition characterised by recurrent, unprovoked seizures.(73) Epilepsy is a heterogenous group of disorders that have been classified by the International League Against Epilepsy (ILAE).(74) Approximately 60% of patients diagnosed with epilepsy have localisation related (focal, local, partial) epilepsy, where generalised epilepsy accounts for approximately 40% of cases.(75) Epilepsy is a common disorder worldwide, affecting 50 per 100,000 people,(76) with an estimated 400,000 individuals in the

UK, and responsible for over 1,000 deaths per year as a result of poor management.(77)

4.2.2 Current Management

Although a common condition, there are relatively few epilepsy specialists within the NHS, with many people with epilepsy being diagnosed and treated by non-specialists in both primary and secondary care. Consequently, there is data to support the claim that the management of this condition can often be sub-optimal.(78) Areas of specific concern include initial diagnosis and drug management.

4.2.3 Licensing Regulations

In the era of evidence-based practice, which is increasingly governing the practice of physicians, the choice of which AED to prescribe should be based upon the results of comparative randomised controlled trials (RCTs), or systematic reviews, however, at present there is insufficient guidance to assist clinicians on which AED to preferentially prescribe for patients with refractory epilepsy.(79) This is the consequence of pharmaceutical companies creating trial designs in order to comply with regulatory criteria (rather than a more enhanced clinically relevant design) as this is the minimum required to attain a Marketing Authorisation.

In the first instance, new AEDs are tested in patients with refractory partial epilepsy as add-on treatment in randomised, placebo-controlled trials. Few trials are subsequently conducted to compare such agents with traditional AEDs or indeed another new AED. Consequently, the selection of an AED for a patient by their physician remains based on an empirical assessment of the probability of achieving seizure freedom and associated side-effect profiles, rather than on the rational application of a treatment that corrects a specific functional or biochemical abnormality.(80) Based on this, there are currently no reliable tools to predict which AED, among the many known to be active against a given seizure type, will control the seizures in an individual patient; AED therapy is thus based mainly on a trial-and-error approach.(81)

4.2.4 Current Evidence and Gaps in the Literature

For patients with *newly-diagnosed* partial and generalised tonic-clonic seizures, a variety of comparative, randomised, double-blind trials suggest similar efficacy [time to treatment failure, 12 month remission, or time to first seizure] for the traditional AEDs; phenytoin, carbamazepine, and sodium valproate, as for the newer AEDs; lamotrigine, and oxcarbazepine.(82-88)

With regards to *refractory partial-onset epilepsy*,(89) have previously conducted a meta-analysis, showing that all of the newer AEDs are clearly more effective than placebo; however, with overlapping confidence intervals, the authors were unable to conclude which one was the most effective. Tolerability was also considered within this analysis, where the same conclusion was reached. As this analysis was performed over a decade ago, with the inclusion of unpublished studies and without some of the AEDs licensed in the UK at the present day, its conclusions are now outdated.

In 2000, the ILAE published an evidence-based guideline to assist clinicians with the treatment of epilepsy, however, a hierarchy was not presented within the final document. This was due in part to a lack of published head-to-head randomised-controlled trials.

Following a systematic review of published and unpublished data by NICE, a clinical guidance document (CG20) was published, providing guidance to physicians on which AEDs to prescribe at what stage and for which epilepsy classifications.(90) This guidance, however, still requires the physician to make a choice between up to five recommended AEDs per classification. An updated version has been published (CG137) however the limitation in their conclusion remains.(91)

4.2.5 Issues with Current Literature

Organisations actively involved in the production of guidelines on various aspects of epilepsy therapy include the ad-hoc

subcommittees of the American Academy of Neurology (AAN), and the American Epilepsy Society (AES); as well as the Scottish Intercollegiate Guidelines Network (SIGN); the U.K. National Institute of Clinical Excellence (NICE); and the International League Against Epilepsy (ILAE).

Of the reviews conducted to date, none of the above have conclusively stated which of the AEDs available should be prescribed preferentially. Subsequently, the anticipated guidelines produced were not able to fully inform clinical practice as they did not present their findings in the form of a treatment algorithm. The major usefulness of this work however is that it highlighted the paucity of well-designed trials in this area of medicine where inadequate statistical inferences are often made to power the data to deliver a clinically meaningful result.(92)

4.3 Rationale for this Research Project

A recent review conducted by the National Institute for Health Research (NIHR, 2008) concluded that although a number of new AEDs have been licensed over the past decade, for several reasons, the trial data upon which these licenses were based have failed to inform clinical practice or policy.(93) A recent Health Technology Appraisal published by NICE also indicated that there is a need for more direct comparative data of newer *versus* newer, and newer *versus* older AEDs within clinical trials, with consideration given to different treatment sequences within both monotherapy and adjunctive therapy.(94)

Various guidelines have been published over the past decade; however, physicians still continue to make individual prescribing decisions.(95) Consequently, significant variability in prescribing continues to be common place in clinical practice. This situation is perpetuated by physicians being presented with an increasing choice of new AEDs to prescribe for patients with refractory epilepsy, as data suggests their superiority over the traditional AEDs.(96)

Although RCTs are the gold standard of research methodology, the conduct of a large trial is very time consuming and expensive, hence the reluctance in uptake from the epilepsy research community. Decisions about the utility of an intervention or the validity of a hypothesis should be taken with caution based on the results of small RCTs, as results typically vary from one study to the next. Systematic reviews are increasingly conducted to summarise the available published data.

4.4 Collaboration

All work undertaken within this chapter of the thesis was undertaken by me. Systematic review support was provided by Dr Grosso in the form of a second independent check against the proforma. Advice and tuition on the use of WinBUGS was provided by Mr Wonderling with all analyses performed by me.

4.5 Protocol

4.5.1 Systematic Review

Relevant randomised trials were searched for within the Cochrane Central Register of Controlled Trials (Cochrane Library 2009, issue 2), which contains the Epilepsy Group's specialised register, Medline (1950–March 2009), Embase (1980–March 2009) and Current Contents Connect [part of the ISI Web of Knowledge] (1998–March 2009). The search terms and limits are provided below (see section 4.5.4).

In addition to the database search strategy, the reference lists of identified manuscripts were manually hand-searched to identify additional relevant studies. To formulate and ensure optimal reporting of the systematic review and meta-analysis, the established PRISMA [preferred reporting items for systematic reviews and meta-analysis] statement was followed.(97)

4.5.1.1 Inclusion Criteria

Trials were included if they were of randomised, double-blind, placebo-controlled or active-controlled add-on design

(investigating acetazolamide, carbamazepine, clobazam, clonazepam, ethosuximide, gabapentin, lacosamide, lamotrigine, levetiracetam, oxcarbazepine, phenobarbital, phenytoin, pregabalin, primidone, sodium valproate, tiagabine, topiramate, valproic acid, vibabatrin or zonisamide), provided they recruited adult patients (> 18 years) with simple / complex partial seizure with or without secondary generalised tonic-clonic seizures, were of parallel or cross-over design, of at least 8 weeks duration and reported seizure frequency and / or adverse effects as an outcome.

4.5.1.2 Exclusion Criteria

Studies were excluded if they contained a pre-randomisation run-in response-conditional design where patients were allocated treatment only if they showed a pre-determined response or if randomisation was preceded by an open-label period in order to minimise the inclusion of data from an enriched population.

Additional exclusion criteria were: trials which incorporated a surgical intervention; trials which used other therapies which may affect seizure frequency; trials of open-label design; observational studies; conference proceedings; and publications available only in abstract form. Non-English language publications were also excluded (see Figure 13).

The purpose of imposing these specific exclusion criteria was to permit extraction of data from the most robustly conducted studies of a similar design. Eligible studies identified were cross-checked against previous systematic reviews.

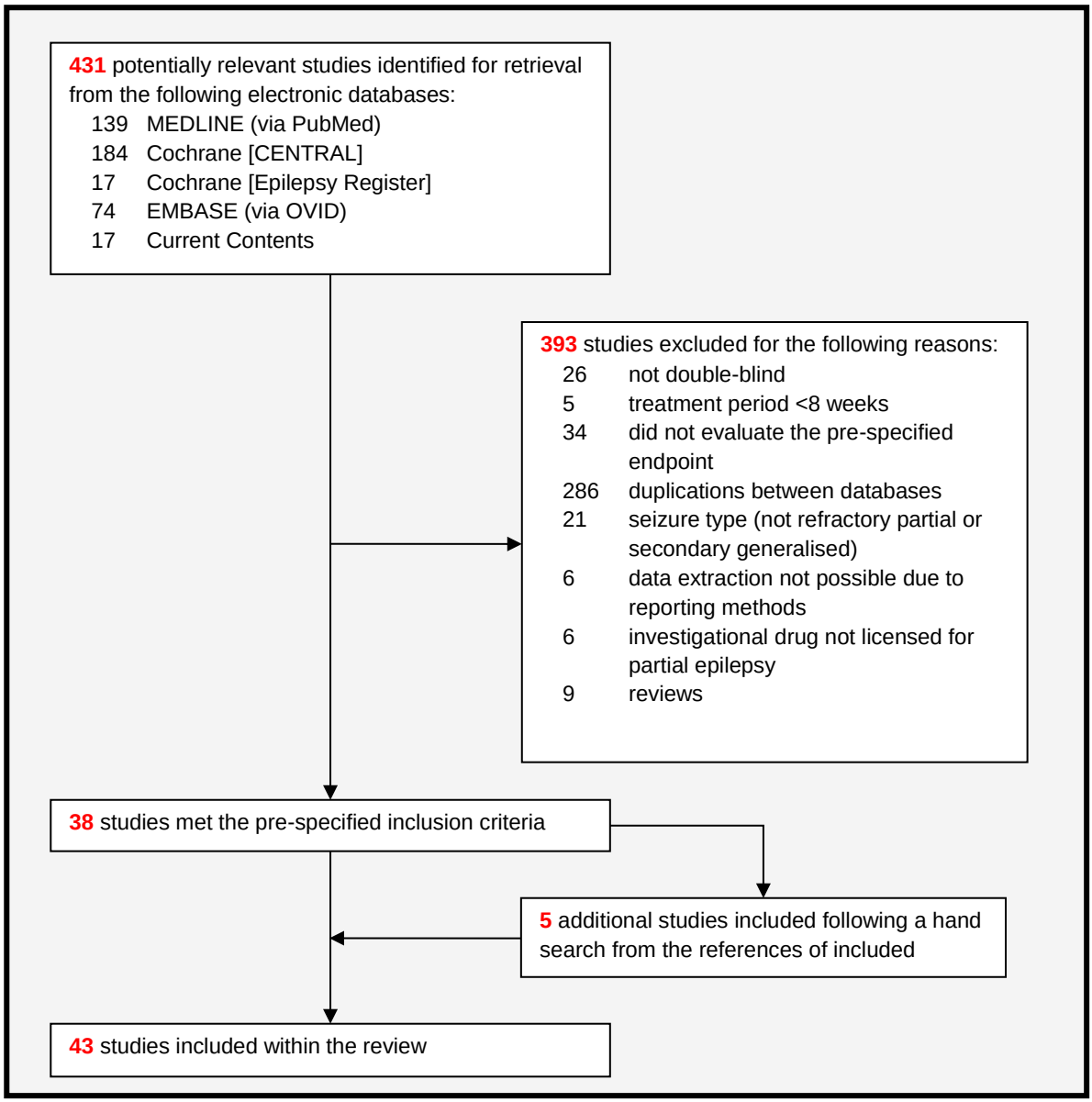


Figure 13: The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of studies identified from the systematic review for inclusion in the network meta-analysis for refractory epilepsy

4.5.2 Endpoints

Two primary endpoints were included as part of this review:

1) The efficacy endpoint, specified *a priori*, was responder rate defined as a 50% reduction in seizure rate from baseline, as recommended by the European Medicines Agency (EMA) for investigations of medicinal products for the management of epilepsy. For studies that did not report their results in this manner, we calculated the proportion of patients with a 50% reduction in seizure rate through analysis of seizure diary rate at end of study minus baseline seizure rate.

2) The tolerability endpoint, specified *a priori*, was the incidence of premature withdrawal from treatment due to drug-related adverse events. Where trials included groups randomized to different doses of the same drug, we selected the group receiving the dose closest to that in established use based on the clinical experience of senior members of the research project.

4.5.3 Quality Assessment and Data Collection

Data from systematic reviews or previous meta-analyses were not used to enable the collection of data from original sources; however, any such publications identified served as a comparator to ensure that all relevant studies had been included within this review. Secondary searches were conducted from the reference lists of manuscripts identified. To minimise bias, a second investigator (AMG) was included at this stage to review the references and abstracts retrieved by the search and select potentially relevant publications against the pre-specified inclusion and exclusion criteria. Important clinical and methodological study characteristics were extracted onto a standard form, checked and recorded to ensure consistency. This included:

1) Characteristics of trial participants (including age, prior therapy, and seizure type)
2) Type of intervention (including type, dose, and duration)
3) Type of outcome measure.

Any discrepancies or lack of agreement between the two reviewers were referred to a third independent investigator (ADH) for arbitration. An assessment of risk of bias (using established criteria)(27) was also undertaken. All analyses were based on *intention to treat* data. For any trials that reported data using a *per-protocol* analysis, *intention to treat* values were calculated.

4.5.4 Search Terms

See Table 2 and Table 3 for a list of the search terms and limits applied when conducting the database search.

Table 2: Search strategy used to determine eligible published trials

	Search term
01.	Epilepsy/
02.	Epilepsy.tw
03.	Seizures/
04.	Seizure$.tw
05.	Convulsions/
06.	Convulsion$.tw
07.	Or/1-6
08.	Acetazolamide/
09.	Carbamazepine/
10.	Clobazam/
11.	Clonazepam/
12.	Ethosuximide/
13.	Gabapentin/
14.	Lacosamide/
15.	Lamotrigine/
16.	Levetiracetam/
17.	Oxcarbazepine/
18.	Phenobarbital/
19.	Phenytoin/
20.	Pregabalin/
21.	Primidone/
22.	Sodium valproate/
23.	Tiagabine/
24.	Topiramate/
25.	Valproic acid/
26.	Vigabatrin/
27.	Zonisamide/
28.	Or/8-27
29.	Monotherapy.tw
30.	Adjunctive.tw
31.	29 or 30
32.	Randomized controlled trials/
33.	Randomized-controlled-trial.pt
34.	Or/32-33
35.	Treatment outcome.tw
36.	Tolerability/
37.	Safety/
38.	Side effects/

39.	Adverse effects/
40.	Adverse events/
41.	Or/35-40
42.	7 and 28 and 31 and 34 and 41

Table 3: Limits applied to each database as part of the systematic review

Database	Limits applied
PubMed	Humans; Randomised controlled trial; English; all adult: 19+ years
Cochrane	Central Register of Controlled Trials
Embase	Full text; Human; English language ; Article OR Erratum; Adult <18 to 64 years> OR aged <65 years+
Current Contents	English language; Article [document type]; Randomised Controlled Trial

4.5.5 Quantitative (Statistical) Analysis – Frequentist Meta-Analysis

Statistical analyses were performed in two stages. For the first stage, a random-effects 'frequentist' meta-analysis(38) was conducted, for placebo-controlled trials of AEDs for refractory epilepsy. The endpoint of interest was the calculation of a pooled odds ratio (OR) with corresponding 95% confidence interval (CI) for both the efficacy and tolerability endpoints. To evaluate heterogeneity of the effect estimates the Cochran Q (Chi-squared), Higgins I-squared and tau-squared statistics were generated.(48)

To explore potential inconsistency and small study bias further l'Abbé and Funnel plots, respectively, were generated.

To generate the Forest plots and Funnel plots, Review Manager (RevMan) version 5.0 (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2008) was used. To generate the l'Abbé plots, StatsDirect® 2.7.7 (Altrincham, Cheshire, UK) was used.

4.5.6 Statistical Analysis – Bayesian Network Meta-Analysis

For the second stage, the relative effectiveness and tolerability profile of each AED was analysed using an extension of the multivariate Bayesian hierarchical random-effects model for mixed multiple treatment comparisons using minimally informative prior distributions.(31, 98-100) The model included random-effects at the level of trials (see Figure 14 and Figure 15) which allowed the estimation of the variance of treatment effects between trials. Due to

the non-discrete methodology used within trials investigating antiepileptic drugs (i.e. dose escalation over the active period) the network meta-analysis was performed using only one evidence network i.e. the different agents (at the most appropriate dose) were treated as separate nodes.

The control event rate used within the model was 0.16 for efficacy and 0.05 for tolerability. These were calculated as the mean across the studies which met the inclusion criteria as a means of determining the placebo-corrected response across the treatment network. The control event rate is synonymous with the minimally-informative prior.

As number needed to treat (NNT) and number needed to harm (NNH) have increasingly been noted as being more clinically useful parameters in highlighting treatment effects, these summary estimates were calculated from the relative risk (RR) estimates generated by WinBUGS using one of the following equations: 1/(1 - RR) x CER if the RR was <1 or 1/(RR - 1) x CER if the RR was >1.(22) The 95% CrI values for NNT and NNH were estimated using the 95% CrI estimated for the RR using the same formulae above. A treatment hierarchy was also developed on the basis that the (posterior) probability that each treatment is the best (using the 95% CrI for the relative rank of each AED).

Figure 14: WinBUGS model for the AED network meta-analysis (efficacy endpoint)

Random effect model: Includes correlation structure for 3-arm trials

```
➔
model{
sw[1] <- 0
for(i in 1:84)  {
      logit(p[i])<-mu[s[i]]+ delta[i]  * (1-equals(t[i],b[i]))                    #  model
      r[i]~dbin(p[i],n[i])                                                       #  binomial likelihood
       delta[i] ~ dnorm(md[i],taud[i])I(-5,5)                                    # trial-specific LOR distributions
       taud[i] <- tau * (1 + equals(m[i],3) /3)                                  # precisions of LOR distributions
      md[i] <- d[t[i]] - d[b[i]]  +  equals(m[i],3) * sw[i]                      # means of LOR distributions

#Deviance residuals for data i
      rhat[i] <- p[i] * n[i]
      dev[i] <- 2 * (r[i] * (log(r[i])-log(rhat[i]))  +  (n[i]-r[i]) * (log(n[i]-r[i]) - log(n[i]-rhat[i])))

     }

sumdev <- sum(dev[])                                                            # Calculate residual deviance


for (i in 2:84)  {   sw[i] <- (delta[i-1] -  d[t[i-1]] + d[b[i-1]])/2}13         # adjustment for 3-arm trials

for(j in 1:42){ mu[j]~dnorm(0,.0001) }                                          # vague priors for NS trial baselines

d[1]<-0
for (k in 2:12)  {d[k] ~ dnorm(0,.0001) }                                       # vague priors for basic parameters

sd~dunif(0,2)                                  #  vague prior for random effects standard deviation
tau<-1/pow(sd,2)

rr[1]<-1
for (k in 2:12)  {logit(v[k])<-logit(0.16 )+d[k]
rr[k]<-v[k]/0.16  }                            # calculate relative risk

for (k in 1:12)
{ ar[k]<-0.16*rr[k]
arr[k]<-0.16-ar[k]
nnt[k]<--1/arr[k]  }    # calculate NNT

# Ranking and prob{treatment k is best}
 for (k in 1:12) {
            rk[k]<-rank(rr[],k)
best[k]<-equals(rank(rr[],k),1)}

# Pairwise ORs
for (c in 1:(12-1))
        {  for (k in (c+1):12)
             {  lor[c,k] <- d[k] - d[c]
                log(or[c,k]) <- lor[c,k]
             }
         }
}
➔
```

Figure 15: WinBUGS model for the AED network meta-analysis (tolerability endpoint)

Random effect model: Includes correlation structure for 3-arm trials

```
➔
model{
sw[1] <- 0
for(i in 1:80)  {
      logit(p[i])<-mu[s[i]]+ delta[i]  * (1-equals(t[i],b[i]))                    #  model
      r[i]~dbin(p[i],n[i])                                                       #  binomial likelihood
       delta[i] ~ dnorm(md[i],taud[i])I(-5,5)                                    # trial-specific LOR distributions
       taud[i] <- tau * (1 + equals(m[i],3) /3)                                  # precisions of LOR distributions
      md[i] <- d[t[i]] - d[b[i]]  +  equals(m[i],3) * sw[i]                      # means of LOR distributions

#Deviance residuals for data i
      rhat[i] <- p[i] * n[i]
      dev[i] <- 2 * (r[i] * (log(r[i])-log(rhat[i]))  +  (n[i]-r[i]) * (log(n[i]-r[i]) - log(n[i]-rhat[i])))

     }

sumdev <- sum(dev[])                                                            # Calculate residual deviance

for (i in 2:80)  {   sw[i] <- (delta[i-1] -  d[t[i-1]] + d[b[i-1]])/2}          # adjustment for 3-arm trials

for(j in 1:40){ mu[j]~dnorm(0,.0001) }                                          # vague priors for NS trial baselines

d[1]<-0
for (k in 2:12)  {d[k] ~ dnorm(0,.0001) }                                       # vague priors for basic parameters

sd~dunif(0,2)                                  #  vague prior for random effects standard deviation
tau<-1/pow(sd,2)

rr[1]<-1
for (k in 2:12)  {logit(v[k])<-logit(0.049898392 )+d[k]
rr[k]<-v[k]/0.049898392  }                                                      # calculate relative risk

for (k in 1:12)
{ ar[k]<-0.049898392*rr[k]
arr[k]<-0.049898392-ar[k]
nnh[k]<--1/arr[k]  }    # calculate NNH

# Ranking and prob{treatment k is best}
 for (k in 1:12) {
            rk[k]<-12+1-rank(rr[],k)
best[k]<-equals(12+1-rank(rr[],k),1)}

# Pairwise ORs
for (c in 1:(12-1))
        {  for (k in (c+1):12)
             {  lor[c,k] <- d[k] - d[c]
                log(or[c,k]) <- lor[c,k]
             }
         }
}
➔
```

4.6 Findings

Out of 431 potentially eligible reports, a total of 43 studies, describing 11 AEDs and including 8,546 patients with refractory epilepsy, met the inclusion criteria and were included within the network meta-analysis.(101-112) All studies were published as full journal articles.

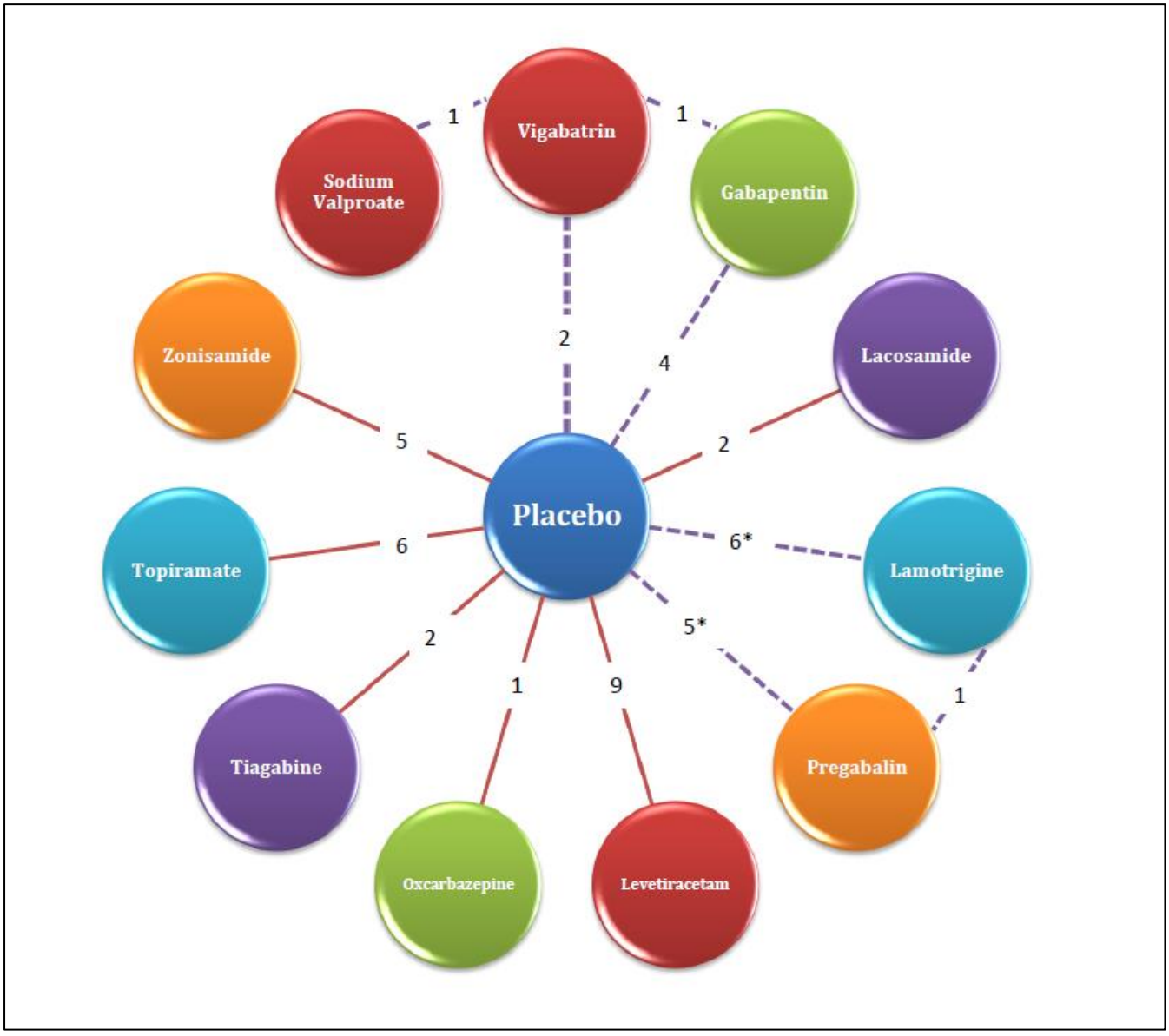


Figure 16: AEDs included within the network meta-analysis.

Each AED represents a node within the star-shaped network. The lines between the nodes represent direct comparative data, where the number along the line indicates the number of studies for that particular link within the network. The purple (dotted) line represents a loop of direct comparative data which allows mixed treatment comparison (*includes the 3-arm study)

4.6.1 Study Characteristics

The main characteristics of the studies which met the pre-specified inclusion criteria and subsequently included within the analysis are summaries in Table 4 and Table 5.

Of the 43 studies, 40 were placebo-controlled, two were active-controlled (sodium valproate vs. vigabatrin, and gabapentin vs. vigabatrin), and one was a three-arm study (pregabalin vs. lamotrigine vs. placebo).(112) Figure 16 displays the treatment network graphically.

The mean number of adjunctive baseline medications was between 1 and 3. The duration of maintenance therapy ranged from 8 to 24 weeks. None of the studies investigating acetazolamide, carbamazepine, clobazam, clonazepam, ethosuximide, phenobarbital, phenytoin or primidone fulfilled the inclusion criteria. Reasons for exclusion were a lack of double-blinding or trial duration of less than 8 weeks. Twenty of the 43 studies (6,212 patients) investigated the AED against placebo at more than one dose. Of 65 possible pair-wise comparisons between the available interventions, including placebo, only 13 were actually reported and in all but three the comparator was placebo.

Table 4: Characteristics of studies included within the AED meta-analysis and network meta-analysis

Study	Study design (weeks of treatment*)	Daily treatment (number of subjects allocated)	Male:Female	Age range (years**)	Type of epilepsy	Number of concomitant AEDs (specified where stated)
Placebo-controlled studies						
Gabapentin						
UK Gabapentin Study Group (1990)(101)	Parallel (14)	**GBP** 1200mg (61), **placebo** (66)	53:74	14-73	SPS/CPS ± GTCS	≤ 2
Sivenius et al (1991)(102)	Parallel (12)	**GBP** 900mg (18), **GBP** 1200mg (9), **placebo** (18)	20:23	16-59	SPS/CPS ± GTCS	≤ 2 (CBZ, CLZ, VPA, PHT)
Anhut et al (1994)(103)	Parallel (12)	**GBP** 900mg (111), **GBP** 1200mg (52), **placebo** (109)	154:120	12-67	SPS/CPS ± GTCS	≤ 2
Yamauchi et al (2006)(104)	Parallel (12)	**GBP** 1200mg (86), **GBP** 1800mg (41), **placebo** (82)	101:108	≥ 16	SPS/CPS	≤ 2
Lacosamide						
Ben-Menachem et al (2007)(105)	Parallel (12)	**LCS** 600mg (106), **LCS** 400mg (108), **LCS** 200mg (107), **placebo** (97)	192:226	18-68	SPS/CPS ± GTCS	≤ 2
Halasz et al (2009)(106)	Parallel (12)	**LCS** 400mg (159), **LCS** 200mg (163), **placebo** (163)	250:235	16-70	SPS/CPS ± GTCS	≥ 2
Lamotrigine						
Loiseau et al (1990)(107)	Crossover (2x8)	**LTG** (23), **placebo** (23) [different doses]	12:11	21-54	SPS/CPS ± GTCS	≤ 2
Matsuo et al (1993)(109)	Crossover (2x12)	**LTG** (126), **placebo** (67) [different doses]	60:133	18-63	SPS/CPS ± GTCS	≤ 3
Messenheimer et al (1994)(108)	Crossover (2x12)	**LTG** (88), **placebo** (98) [different doses]	46:52	18-64	SPS/CPS ± GTCS	≤ 3
Biton et al (2005)(110)	Parallel (12)	**LTG** (58), **placebo** (59) [different doses]	62:55	2-55	GTCS	≤ 2 (PHT, PHB, CBZ, PRM, VPA)
Naritoku et al (2007)(111)	Parallel (12)	**LTG** (118), **placebo** (121) [different doses]	117:119	≥ 12	SPS/CPS ± GTCS	≤ 2
Baulac et al (2010)(112)	Parallel (17)	**LTG** (141), **placebo** (141) [different doses]	132:149	16-67	SPS/CPS	≤ 3
Levetiracetam						
Cereghino et al (2000)(113)	Parallel (14)	**LEV** 1000mg (98), **LEV** 3000mg (101), **placebo** (95)	178:116	16-75	SPS/CPS ± GTCS	≥ 2
Shorvon et al (2000)(114)	Parallel (12)	**LEV** 1000mg (106), **LEV** 2000mg (106), **placebo** (112)	157:167	16-65	SPS/CPS ± GTCS	≤ 2
Betts et al (2000)(115)	Parallel (24)	**LEV** 2000mg (42), **LEV** 4000mg (38), **placebo** (39)	73:46	16-70	SPS/CPS ± GTCS	≤ 3
Ben-Menachem et al (2000)(116)	Parallel (12)	**LEV** 3000mg (181), **placebo** (105)	137:149	16-70	CPS	1
Boon et al (2002)(117)	Crossover (2x12)	**LEV** 2000mg (202), **placebo** (200)	157:167	16-65	SPS/CPS ± GTCS	≥ 3
Tsai et al (2006)(118)	Parallel (12)	**LEV** 2000mg (47), **placebo** (47)	42:52	16-60	SPS/CPS ± GTCS	≤ 3
Peltola et al (2009)(119)	Parallel (12)	**LEV** 1000mg (79), **placebo** (79)	99:59	12-68	SPS/CPS ± GTCS	≤ 3
Zhou et al (2008)(120)	Parallel (12)	**LEV** 3000mg (13), **placebo** (11)	13:11	16-70	SPS/CPS ± GTCS	≤ 2

Study	Study design (weeks of treatment*)	Daily treatment (number of subjects allocated)	Male:Female	Age range (years**)	Type of epilepsy	Number of concomitant AEDs (specified where stated)
Xiao et al (2009)(121)	Parallel (12)	**LEV** 3000mg (28), **placebo** (28)	24:36	17-60	SPS/CPS ± GTCS	≤ 2 (PHT, CBZ, PHB, PRM, VPA, TPM, GBP, LMG)
Oxcarbazepine						
Barcs et al (2000)(122)	Parallel (24)	**OXC** 600mg (168), **OXC** 1200mg (177), **OXC** 2400mg (174), **placebo** (173)	341:351	15-65	SPS/CPS ± GTCS	≤ 3
Pregabalin						
French et al (2003)(123)	Parallel (12)	**PRB** 50mg (88), **PRB** 150mg (86), **PRB** 300mg (90), **PRB** 600mg (89), **placebo** (100)	218:235	12-75	SPS/CPS ± GTCS	≤ 3
Arroya et al (2004)(124)	Parallel (12)	**PRB** 150mg (99), **PRB** 600mg (92), **placebo** (96)	145:142	17-73	SPS/CPS ± GTCS	≤ 3
Beydoun et al (2005)(125)	Parallel (12)	**PRB** 200mg TDS (111), **PRB** 300mg BD (104), **placebo** (98)	156:156	17-82	SPS/CPS ± GTCS	≤ 4
Lee et al (2009)(126)	Parallel (12)	**PRB** 600mg (119), **placebo** (59)	86:92	≥ 18	SPS/CPS ± GTCS	≤ 4 (CBZ, VPA, TPM, LTG, PHB, OXC)
Tiagabine						
Uthman et al (1998)(127)	Parallel (20)	**TGB** 16mg (61), **TGB** 32mg (88), **TGB** 56mg (57), **placebo** (91)	58:42	12-77	SPS/CPS ± GTCS	≤ 3 (PHT, CBZ, PHB, PRM)
Kalviainen et al (1998)(128)	Parallel (22)	**TGB** 30mg (77), **placebo** (77)	90:64	16-75	SPS/CPS ± GTCS	≤ 6 (CBZ, CLZ, PHT, VPA, VGB)
Topiramate						
Tassinari et al (1996)(129)	Parallel (12)	**TPM** 600mg (30), placebo (30)	24:6	18-65	SPS/CPS ± GTCS	≤ 2
Ben-Menachem et al (1996)(130)	Parallel (13)	**TPM** 800mg (28), placebo (28)	23:5	19-63	SPS/CPS ± GTCS	≤ 2
Faught et al (1996)(131)	Parallel (16)	**TPM** 200mg (45), **TPM** 400mg (45), **TPM** 600mg (46), **placebo** (45)	143:38	19-68	SPS/CPS ± GTCS	≤ 2 (CBZ, PHT)
Privitera et al (1996)(132)	Parallel (18)	**TPM** 600mg (48), **TPM** 800mg (48), **TPM** 1000mg (47), **placebo** (47)	152:38	18-68	SPS/CPS ± GTCS	≤ 2
Sharief et al (1996)(133)	Parallel (11)	**TPM** 400mg (23), **placebo** (24)	40:7	18-65	SPS/CPS ± GTCS	≤ 2 (CBZ, PHT, VPA, PHB, PRD)
Yen et al (2000)(134)	Parallel (14)	**TPM** 300mg (23) , **placebo** (23)	19:27	18-54	SPS/CPS	≥ 4 (CBZ, VPA, LTG, PHT)
Vigabatrin						
French et al (1996)(135)	Parallel (12)	**VGB** 3000mg (93), **placebo** (90)	80:102	18-60	CPS ± GTCS	≤ 2 (CBZ, PHT)
Bruni et al (2000)(136)	Parallel (36)	**VGB** 3000mg [mean] (58), **placebo** (53)	61:50	18-50	CPS ± GTCS	≤ 2
Zonisamide						
Schmidt et al (1993)(137)	Parallel (12)	**ZNS** 500mg [mean] (71), **placebo** (68)	81:58	18-59	CPS	≤ 3 (CBZ, PHT, VPA, PHB, PRM)

Study	Study design (weeks of treatment*)	Daily treatment (number of subjects allocated)	Male:Female	Age range (years**)	Type of epilepsy	Number of concomitant AEDs (specified where stated)
Faught et al (2001)(138)	Crossover (20)	**ZNS** 100mg, 200mg, 400mg (118), **placebo** (85)	104:99	13-68	SPS/CPS ± GTCS	≤ 2 (CBZ, PHT, VPA, PHB, PRM)
Brodie et al (2004)(139)	Parallel (12)	**ZNS** 400mg (73), **placebo** (71)	85:59	18-59	SPS/CPS ± GTCS	≤ 2 (CBZ, PHT, VPA, PHB, PRM)
Sackarelles et al (2004)(140)	Parallel (12)	**ZNS** 500mg [mean] (78), **placebo** (74)	101:51	17-68	CPS ± GTCS	≤ 2 (CBZ, PHT, PHB, PRM)
Brodie et al (2005)(141)	Parallel (24)	**ZNS** 100mg (56), **ZNS** 300mg (55), **ZNS** 500mg (118), **placebo** (120)	232:171	12-77	SPS/CPS ± GTCS	≤ 4 (CBZ, CLB, GBP, LTG, PHB, PHT, TPM, VPA)

*Weeks of treatment refers to the double-blind period consisting of both the titration-to-target and target-stabilisation phases.

**Due to the methodology used in categorising trials within PubMed and Embase, the inclusion criteria of 'all adults 19+ years' and 'adults 18 to 64 years' also included a small number of studies which enrolled both adults and children (see page 121 for further detail)

For those studies where more than one dose of the active agent was compared against placebo, the dose underlined reflects the dose for which data were extracted from the trial to permit comparison.

SPS = simple partial onset seizure; CPS = complex partial onset seizure; GTCS = secondary generalised tonic-clonic seizure

GBP = gabapentin; LCS = lacosamide; LMG = lamotrigine; LEV = levetiracetam; OXC = oxcarbazepine; PRB = pregabalin; VPA = sodium valproate; TGB = tiagabine; TPM = topiramate; VGB = vigabatrin; ZNS = zonisamide

Table 5: Characteristics of additional studies included within the AED network meta-analysis

Study	Study design (weeks of treatment)	Daily treatment (number of subjects allocated)	Male:Female	Age range (years)	Type of epilepsy	Number of concomitant AEDs (specified where stated)
Active-controlled studies						
Vigabatrin and Valproic Acid						
Brodie et al (1999)(84)	Parallel (12)	**VGB** 2000mg to 400mg (108) **VPA** 1000mg to 2000mg (107)	106:109	12-78	SPS/CPS ± GTCS	CBZ
Gabapentin and Vigabatrin						
Lindberger et al (2000)(142)	Parallel (8)	**GBP** 2400mg to 3600mg (50) **VGB** 2000mg to 4000mg (52)	51:51	13-68	SPS/CPS	≤ 2
Lamotrigine and Pregabalin						
Baulac et al (2010)(112)	Parallel (17)	**LTG** 300mg to 400mg (141), **PRB** 300mg to 600mg (152)	155:138	18-82	SPS/CPS	≤ 3

*Weeks of treatment refers to the double-blind period consisting of both the titration-to-target and target-stabilisation phases.

For those studies where more than one dose of the active agent was compared against placebo, the dose underlined reflects the dose for which data were extracted from the trial to permit comparison.

SPS = simple partial onset seizure; CPS = complex partial onset seizure; GTCS = secondary generalised tonic-clonic seizure

GBP = gabapentin; LCS = lacosamide; LMG = lamotrigine; LEV = levetiracetam; OXC = oxcarbazepine; PRB = pregabalin; VPA = sodium valproate; TGB = tiagabine; TPM = topiramate; VGB = vigabatrin; ZNS = zonisamide

4.6.2 Comparison of effect size

4.6.2.1 Frequentist Meta-Analysis

The results of the standard meta-analysis of placebo-controlled trials (stratified by drug) demonstrated that each AED was more efficacious than placebo in reducing seizure events by >50% from baseline (see Figure 17) with an overall OR 3.78 (95% CI 3.14 to 4.55) (see Figure 19). For the efficacy endpoint, there was low to moderate evidence of heterogeneity (tau-squared = 0.15; I-squared = 46%) which was attributable to specific AED drug class (levetiracetam, pregabalin, tiagabine and zonisamide). Similarly, meta-analysis of tolerability indicated a greater overall odds of premature withdrawal due to the development of adverse effects for all AEDs *vs.* placebo (OR 3.27, 95% CI 2.37 to 4.52) (see Figure 18 and Figure 20), with moderate evidence of heterogeneity (tau-squared = 0.45; I-squared = 55%), again attributable to specific AED class (tiagabine, topiramate and zonisamide). Based on these data there was no strong evidence favouring any one particular AED over another on the basis of efficacy, although oxcarbazepine appeared to be the least well tolerated.

No evidence of significant heterogeneity in efficacy or tolerability ($P < 0.05$) was detected in different trials of the same drug following a review of the l'Abbé plots with the possible exception of pregabalin (efficacy analysis Cochran Q = 15.36, p = 0.002). A review of the Funnel plots did not reveal concerns regarding publication bias (see Figure 21 and Figure 22 for the efficacy and tolerability analyses, respectively).

Figure 17: Forest plot (RevMan v5.0) of the odds ratios for **efficacy** (50% responder rate) of randomized controlled trials comparing an AED *vs.* placebo as add-on treatment for refractory epilepsy, respectively.

The black squares represent the odds ratio for individual studies of AED *vs.* placebo and the horizontal line represents the 95% confidence interval of the odds ratio. The black diamond represents the random-effects pooled odds ratio for studies reporting on the same AED where its width represents the 95% confidence intervals. Estimates to the right of the vertical line (i.e. odds ratio >1) are indicative of a statistically significant increase in efficacy, relative to placebo, in patients randomized to the active intervention.

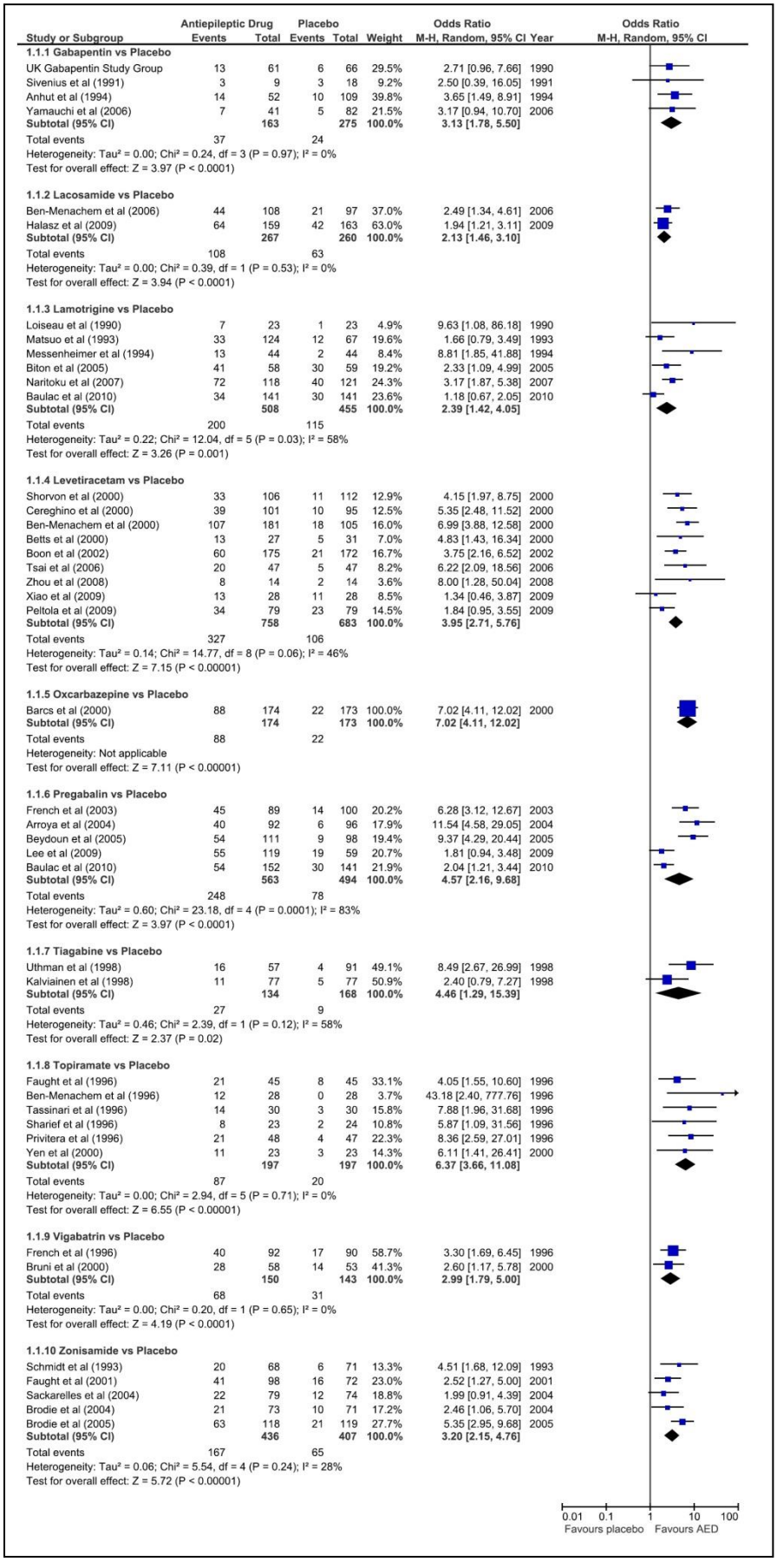


Figure 17: (see above for description)

Figure 18: Forest plot (RevMan v5.0) of the odds ratios for **tolerability** (withdrawal from treatment due to an intolerable adverse event) of randomized controlled trials comparing an AED *vs.* placebo as add-on treatment for refractory epilepsy, respectively.

The black squares represent the odds ratio for individual studies of AED *vs.* placebo and the horizontal line represents the 95% confidence interval of the odds ratio. The black diamond represents the random-effects pooled odds ratio for studies reporting on the same AED where its width represents the 95% confidence intervals. Estimates to the right of the vertical line (i.e. odds ratio >1) are indicative of a statistically significant increase in withdrawal rate, relative to placebo, in patients randomized to the active intervention.

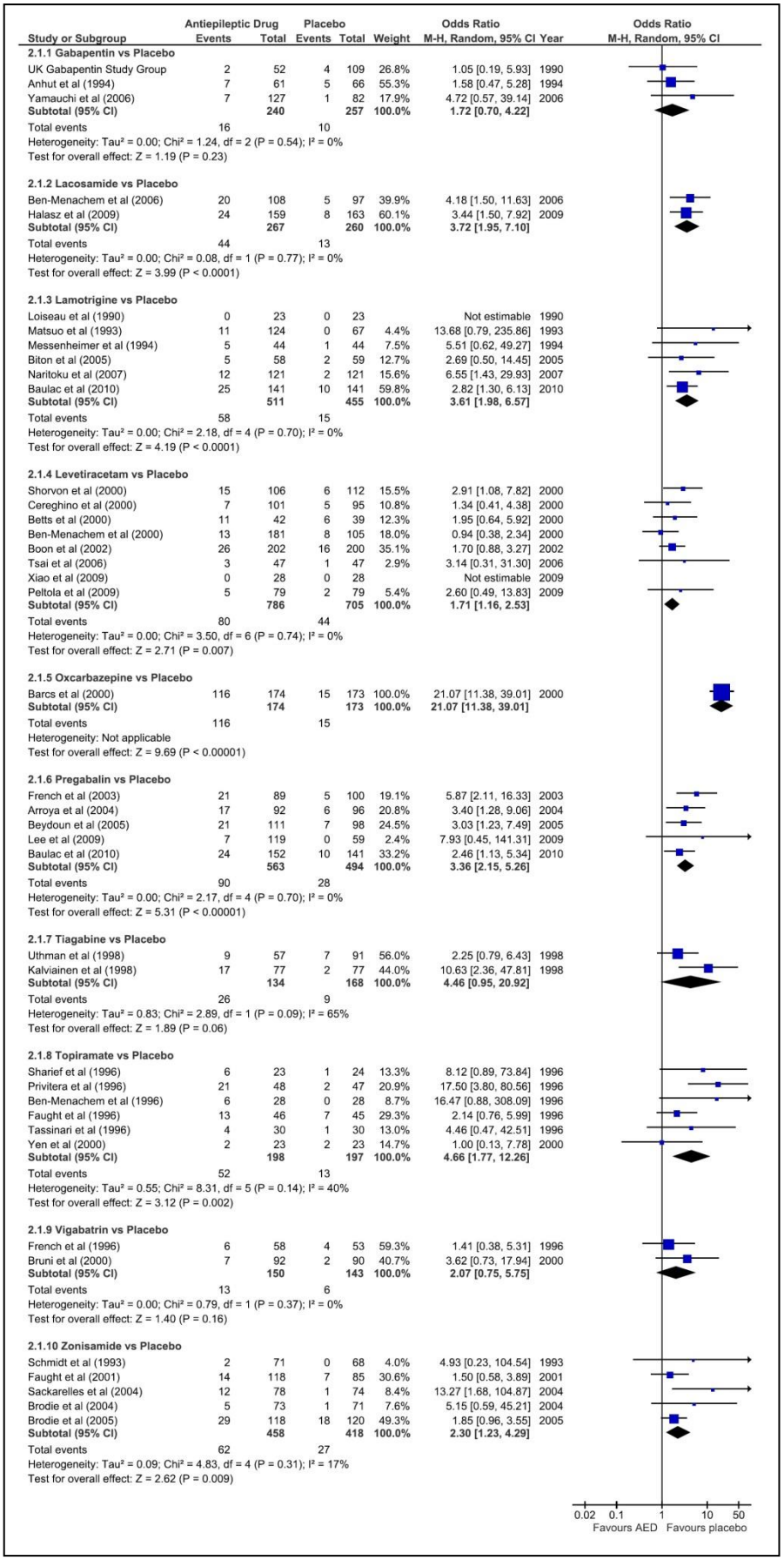


Figure 18: (see above for description)

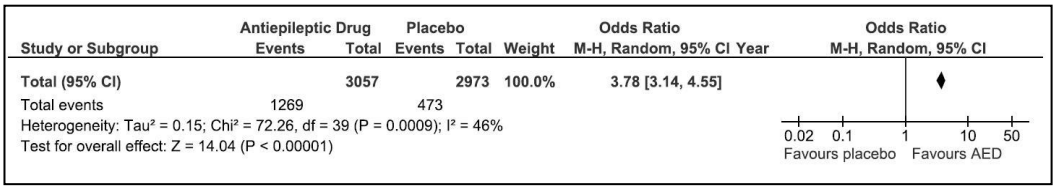


Figure 19: Summary Forest plot (RevMan v5.0) of the odds ratio for **efficacy** (50% responder rate) of randomized controlled trials comparing an AED vs. placebo as add-on treatment for refractory epilepsy

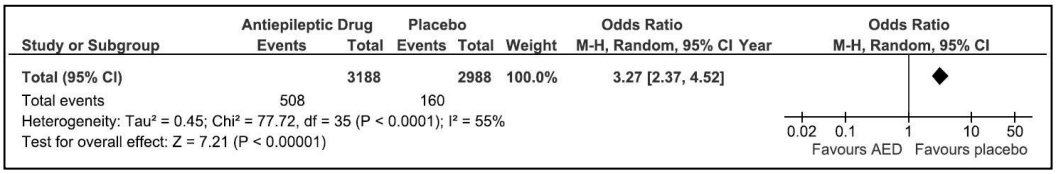


Figure 20: Summary Forest plot (RevMan v5.0) of the odds ratio for **tolerability** (withdrawal from treatment due to an intolerable adverse event) of randomized controlled trials comparing an AED vs. placebo as add-on treatment for refractory epilepsy

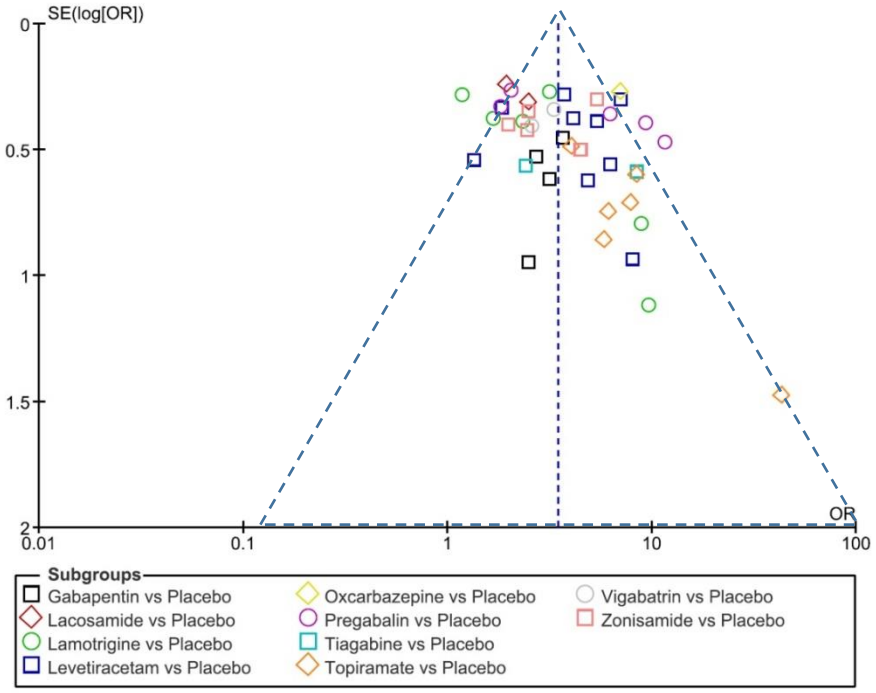


Figure 21: Funnel plot (RevMan v5.0) of randomised controlled trials comparing each AED versus placebo as add-on treatment for refractory epilepsy for the **efficacy** endpoint

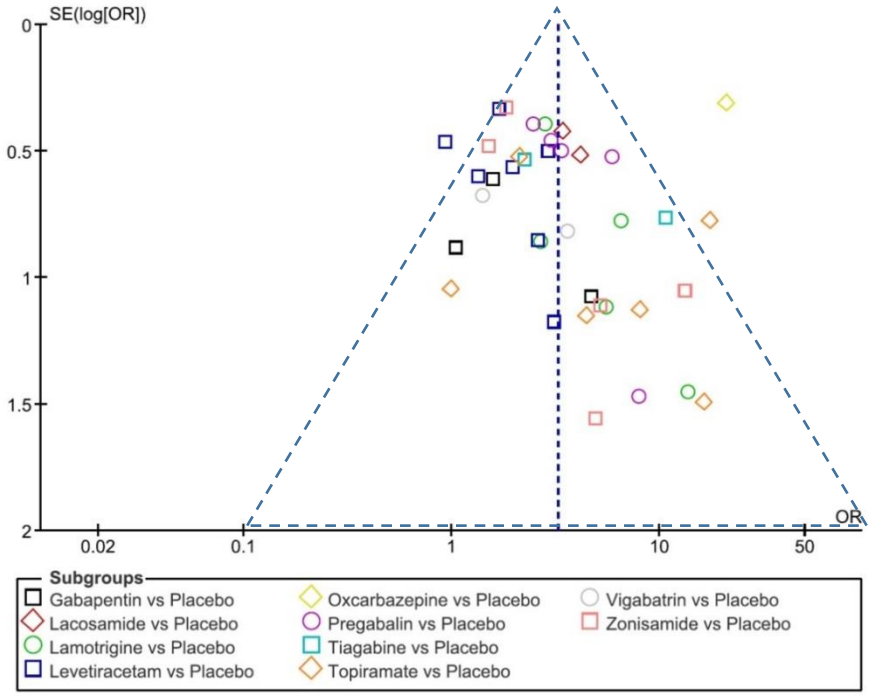


Figure 22: Funnel plot (RevMan v5.0) of randomised controlled trials comparing each AED versus placebo as add-on treatment for refractory epilepsy for the **tolerability** endpoint

4.6.2.2 Bayesian Network Meta-Analysis: efficacy, tolerability and prescribing hierarchy

The median OR estimates, for the efficacy and tolerability endpoints, generated by the Bayesian random-effects network meta-analysis are given in Table 6. As a minimally-informative prior was used, these closely resemble those generated from the Frequentist random-effects meta-analysis (see Table 7 and Table 8 for the efficacy and tolerability endpoints, respectively).

No evidence of significant inconsistency (using Bucher's test) was detected between directly-observed and inferred-treatment effects within the loop identified in Figure 16 for either efficacy (P = 0.26) or tolerability (P = 0.22).

For both outcomes, the model showed reasonable goodness of fit to the data (number of data points) as determined by the posterior mean of the residual deviance [efficacy = 86.0 (84), 95% CrI 62.8, 115.9, P = 0.419; tolerability = 77.6 (80), 95% CrI 57.9, 100.5, P = 0.555]. A visual assessment of the trace plots (history) and time series (density) plots also did not reveal cause for concern regarding inconsistency.

In contrast to the frequentist meta-analysis, network meta-analysis allows the ordering of AEDs according to efficacy and tolerability. Based on the odds ratios estimated, there was an approximately two-fold difference in short term efficacy (based on the 50% responder rate), lacosamide being the least and topiramate the most efficacious at the doses evaluated (see Figure 23). There was an approximately five-fold difference in short term tolerability with valproate being the best and oxcarbazepine being the least well tolerated at the doses evaluated (see Figure 24), however, due to the overlapping 95% CrI associated with these ORs the clinical relevance may not be as large or significant as suggested.

As treatment decisions for refractory partial epilepsy may be based on a balance between efficacy and tolerability, NNT and NNH values were calculated for each drug.

A quantitative trade-off analysis found that four agents (valproate, levetiracetam, gabapentin and vigabatrin) demonstrated the best combination of short term efficacy and tolerability (see Figure 25). Though similarly effective, oxcarbazepine was less well tolerated than all other agents. The remaining agents (topiramate, pregabalin, tiagabine, zonisamide, lamotrigine and lacosamide) demonstrated intermediate short term efficacy and tolerability.

4.6.2.3 Cost minimisation analysis

With consideration to the number of possible AED combinations / permutations, and following discussion of the results with a representative from the neurology service at NHNN, it was decided that a formal cost effectiveness analysis would not be appropriate. Based on the logistics of prescribing in clinical practice, such as the strategy of AED prescribing incorporating patient specific factors (e.g. pregnancy), and the complex prescribing history of AEDs over the last 20 years, such analysis would contain too many confounding factors and override its merits. It was thus agreed that the trade-off analysis should include acquisition cost alongside efficacy and tolerability (see Table 9).

Table 6: Trade-off analysis – comparative efficacy and tolerability of the antiepileptic drugs (AED).

Results are presented as the odds ratios (OR) [and their corresponding 95% credible interval (CrI) values] for the column-defining agent compared with the OR for the row-defining agent. For efficacy, ORs higher than 1 favour the column-defining treatment (e.g. the OR for efficacy for LEV compared with VPA is 1.09). For tolerability, ORs less than 1 favour the row-defining treatment. Results are statistically significant where the 95% CrI for the corresponding OR do not cross 1. To obtain ORs for comparison in the opposite direction, reciprocals should be taken (i.e. the OR for efficacy for VPA compared with LEV is 1/1.09 = 0.92)

Efficacy [response rate] (95% CrI)

Comparison AED

Tolerability [withdrawal rate] (95% CrI)

VPA	0.93 (0.36, 2.38)	0.74 (0.23, 2.62)	0.81 (0.24, 3.11)	2.04 (0.59, 8.41)	1.26 (0.29, 8.41)	1.06 (0.33, 3.81)	1.82 (0.43, 8.69)	0.88 (0.25, 3.11)	1.48 (0.44, 5.64)	0.57 (0.16, 2.21)
1.22 (0.41, 3.59)	VGB	0.80 (0.38, 1.77)	0.89 (0.40, 2.16)	2.21 (0.86, 5.96)	1.37 (0.44, 4.87)	1.62 (0.56, 2.56)	1.98 (0.65, 6.51)	0.95 (0.42, 2.14)	1.61 (0.71, 3.95)	0.62 (0.25, 1.64)
1.23 (0.26, 4.77)	0.97 (0.35, 2.46)	GBP	1.13 (0.50, 2.54)	2.72 (1.16, 6.90)	1.70 (0.56, 5.70)	1.48 (0.70, 2.98)	2.48 (0.82, 7.79)	1.19 (0.52, 2.56)	2.07 (0.89, 4.68)	0.80 (0.31, 1.92)
3.98 (0.69, 22.30)	3.17 (0.88, 13.41)	3.38 (0.99, 13.66)	LMG	2.46 (1.09, 5.92)	1.52 (0.53, 4.90)	1.33 (0.68, 2.48)	2.20 (0.75, 6.63)	1.07 (0.49, 2.12)	1.84 (0.86, 3.86)	0.69 (0.29, 1.66)
3.47 (0.60, 15.44)	2.84 (0.77, 8.58)	2.96 (0.91, 8.20)	0.90 (0.21, 2.63)	TPM	0.63 (0.19, 2.09)	0.53 (0.24, 1.09)	0.92 (0.29, 2.87)	0.42 (0.18, 0.93)	0.74 (0.32, 1.69)	0.28 (0.11, 0.69)
2.84 (0.45, 15.02)	2.28 (0.59, 8.34)	2.41 (0.68, 8.80)	0.70 (0.16, 2.66)	0.85 (0.26, 2.54)	TGB	0.86 (0.28, 2.32)	1.48 (0.36, 5.55)	0.69 (0.21, 1.97)	1.21 (0.37, 3.56)	0.46 (0.13, 1.43)
1.09 (0.23, 4.39)	0.88 (0.29, 2.28)	0.93 (0.35, 2.27)	0.29 (0.08, 0.73)	0.32 (0.14, 0.83)	0.38 (0.13, 1.13)	LEV	1.67 (0.63, 4.82)	0.81 (0.43, 1.49)	1.40 (0.73, 2.75)	0.53 (0.25, 1.17)
13.42 (2.20, 66.34)	10.75 (2.66, 37.18)	11.27 (3.21, 38.06)	3.32 (0.77, 12.13)	3.90 (1.20, 12.68)	4.56 (1.20, 17.88)	12.20 (4.21, 35.40)	OXC	0.48 (0.16, 1.32)	0.83 (0.27, 2.43)	0.37 (0.10, 0.98)
2.43 (0.47, 11.54)	1.97 (0.61, 6.02)	2.04 (0.73, 6.35)	0.61 (0.15, 2.06)	0.70 (0.26, 2.01)	0.83 (0.27, 3.12)	2.17 (1.02, 5.51)	0.18 (0.06, 0.64)	ZNS	1.71 (0.86, 3.71)	0.66 (0.29, 1.55)
2.56 (0.52, 12.22)	2.09 (0.66, 6.51)	2.19 (0.80, 6.71)	0.65 (0.18, 2.08)	0.78 (0.29, 2.23)	0.94 (0.29, 3.17)	2.42 (1.00, 5.96)	0.20 (0.07, 0.65)	1.10 (0.40, 2.89)	PBG	0.38 (0.16, 0.90)
2.41 (0.41, 10.82)	1.93 (0.51, 6.06)	2.00 (0.62, 6.17)	0.59 (0.14, 1.96)	0.67 (0.24, 2.04)	0.83 (0.24, 2.83)	2.16 (0.84, 5.40)	0.17 (0.05, 0.61)	0.98 (0.31, 2.61)	0.88 (0.30, 2.51)	LCS

Table 7: Comparison of odds ratio estimates generated from the frequentist random-effects meta-analysis and the Bayesian network meta-analysis for the **efficacy** endpoint

	Frequentist MA	Bayesian Network MA
Intervention	**Mean Odds Ratio (95% CI)**	**Mean Odds Ratio (95% CrI)**
Sodium Valproate	NA	3.83 (1.15 – 11.62)
Vigabatrin	2.99 (1.79 – 5.00)	3.50 (1.76 – 6.65)
Gabapentin	3.13 (1.78 – 5.50)	2.76 (1.50 – 5.22)
Lamotrigine	2.84 (1.77 – 4.54)	3.12 (1.87 – 5.42)
Topiramate	6.37 (3.66 – 11.08)	7.71 (4.02 – 15.36)
Tiagabine	4.46 (1.29 – 15.39)	4.75 (1.88 – 13.55)
Levetiracetam	3.95 (2.71 – 5.76)	4.11 (2.80 – 6.01)
Oxcarbazepine	7.02 (4.11 – 5.76)	6.87 (2.76 – 18.39)
Zonisamide	3.20 (2.15 – 4.76)	3.30 (1.98 – 5.32)
Pregabalin	5.74 (2.44 – 13.50)	5.73 (3.36 – 9.85)
Lacosamide	2.13 (1.46 – 3.10)	2.17 (1.12 – 4.30)

Table 8: Comparison of odds ratio estimates generated from the frequentist random-effects meta-analysis and the Bayesian network meta-analysis for the **tolerability** endpoint

	Frequentist Meta-Analysis	Bayesian Network Meta-Analysis
Intervention	**Mean Odds Ratio (95% CI)**	**Mean Odds Ratio (95% CrI)**
Sodium Valproate	NA	1.65 (0.44 – 7.62)
Vigabatrin	2.07 (0.75 – 5.75)	2.00 (0.88 – 5.43)
Gabapentin	1.72 (0.70 – 4.22)	1.96 (0.87 – 4.51)
Lamotrigine	5.19 (2.02 – 13.37)	6.32 (2.71 – 21.16)
Topiramate	4.66 (1.77 – 12.26)	5.68 (2.54 – 11.49)
Tiagabine	4.46 (0.95 – 20.92)	4.65 (1.82 – 12.49)
Levetiracetam	1.71 (1.16 – 2.53)	1.80 (1.09 – 2.98)
Oxcarbazepine	21.07 (11.38 – 39.01)	21.86 (8.51 – 56.21)
Zonisamide	2.30 (1.23 – 4.29)	3.96 (2.12 – 8.60)
Pregabalin	3.93 (2.27 – 6.79)	4.29 (2.22 – 9.09)
Lacosamide	3.72 (1.95 – 7.10)	3.85 (1.75 – 8.75)

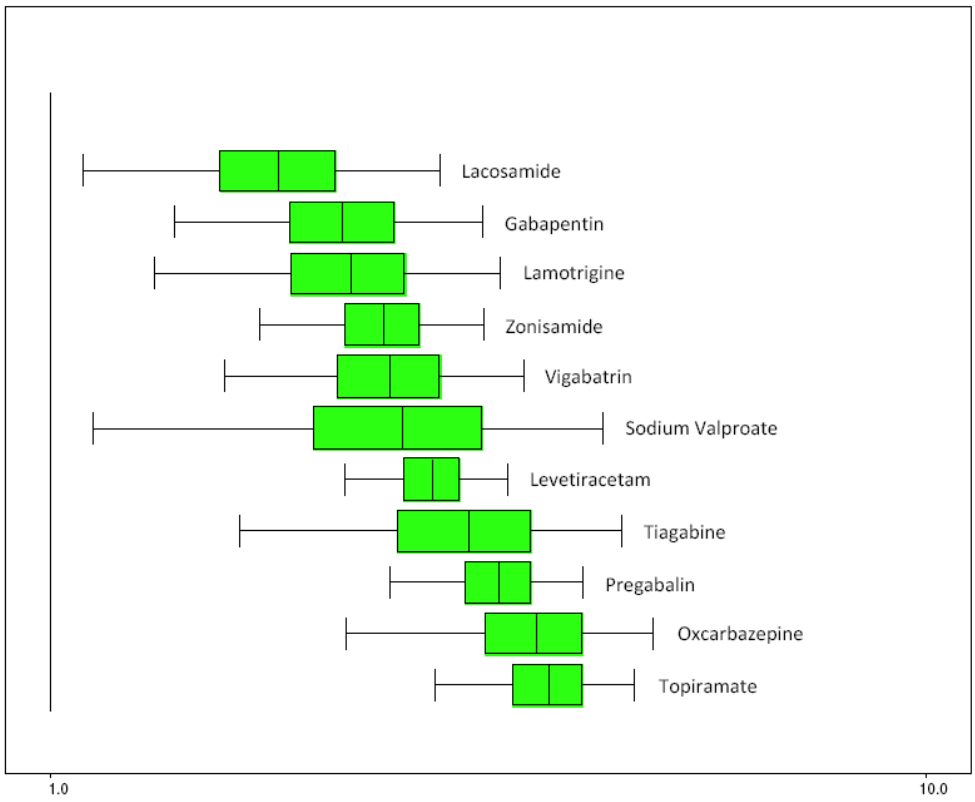


Figure 23: Box plot (WinBUGS) of the risk ratios (RR) of AEDs *vs.* placebo as estimated by the network meta-analysis for the **efficacy** (50% responder rate) endpoint.

The individual plots represent the 95% credible interval (horizontal black line either side of the green box) and the interquartile range (green box, representing where 25% to 75% of the data lies); the vertical black line reflects the median RR. A larger RR is indicative of a greater proportion of patients achieving a 50% reduction in seizure frequency, relative to placebo. Where the 95% credible interval crosses the line of unity, this is indicative of a non-statistically significant difference to placebo. AEDs are listed in descending rank order.

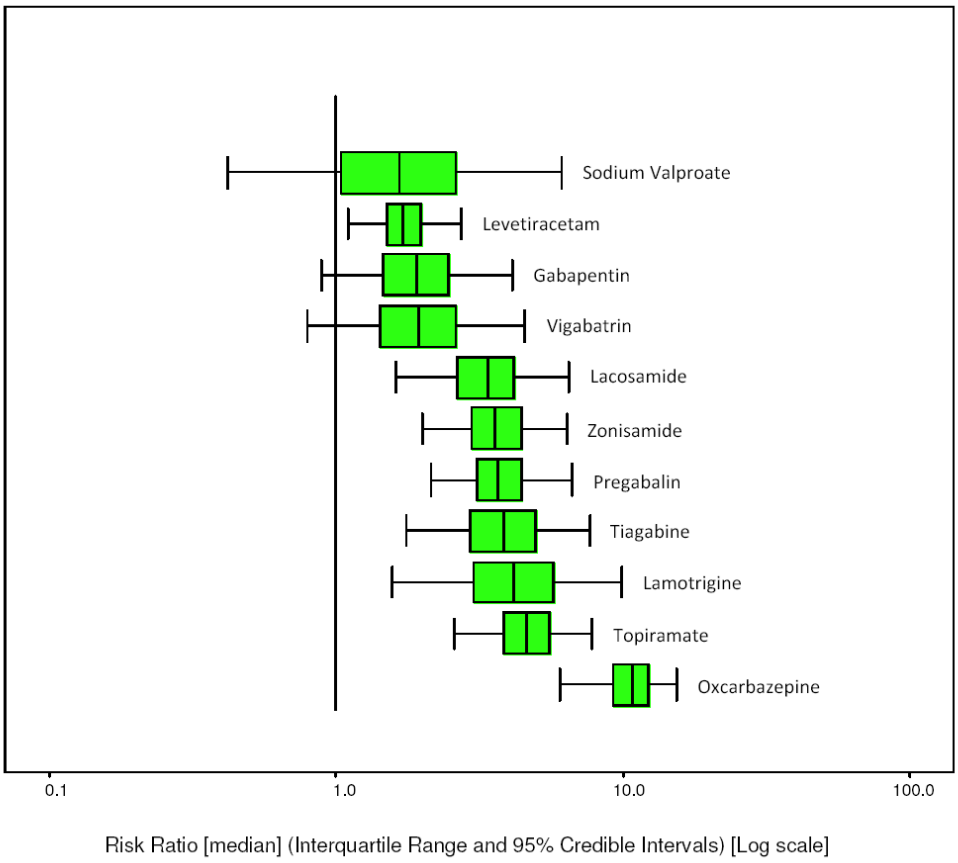


Figure 24: Box plot (WinBUGS) of the risk ratios (RR) of AEDs *vs.* placebo as estimated by the network meta-analysis for the **tolerability** (premature discontinuation) endpoint.

The individual plots represent the 95% credible interval (horizontal black line either side of the green box) and the interquartile range (green box, representing where 25% to 75% of the data lies); the vertical black line reflects the median RR. A larger RR is indicative of a greater proportion of patients withdrawing from treatment prematurely, relative to placebo. Where the 95% credible interval crosses the line of unity, this is indicative of a non-statistically significant difference to placebo. AEDs are listed in descending rank order.

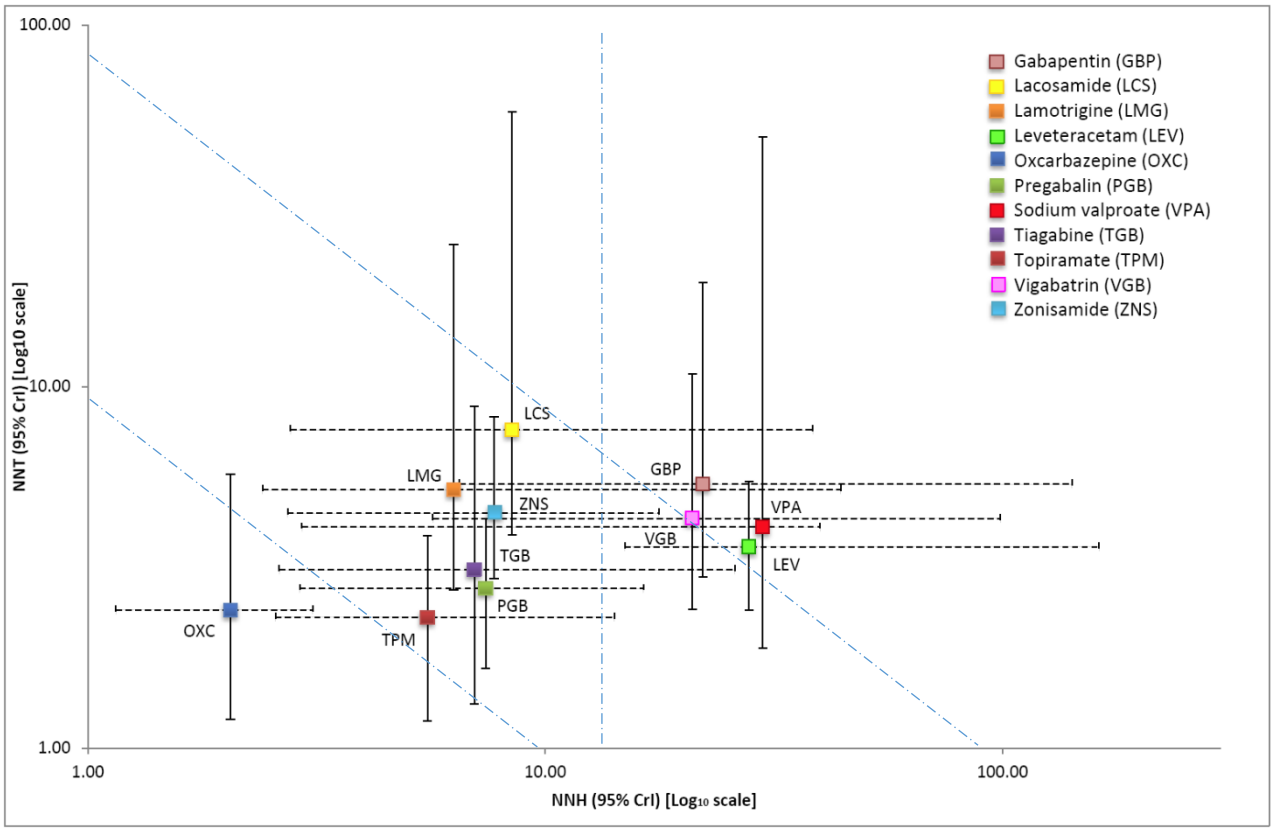


Figure 25: Quantitative trade-off analysis – a comparison of the number needed to treat *vs.* number needed to harm (with 95% credible intervals) for efficacy (50% responder rate) and tolerability (withdrawal rate), endpoints respectively.

Drugs in the lower right of the plot (i.e. low NNT and high NNH) are highly efficacious and well tolerated; conversely, drugs in the upper left of the plot (i.e. high NNT and low NNH) are less effective and poorly tolerated. Note that equal weighting of the efficacy and tolerability endpoints has been applied (i.e. the same emphasis is placed on achieving a 50% reduction in seizure rate in one patient as for incurring a premature withdrawal due to an intolerable adverse effect in another)

Table 9: Trade-off analysis – summary rank of treatment (clinical vs. cost). *AEDs currently protected under patent within the UK

Rank	Efficacy	Tolerability	Cost*
1	Oxcarbazepine	Sodium Valproate	Levetiracetam
2	Topiramate	Levetiracetam	Gabapentin
3	Pregabalin	Gabapentin	Sodium Valproate
4	Tiagabine	Vigabatrin	Lamotrigine
5	Levetiracetam	Lacosamide	Pregabalin
6	Sodium Valproate	Pregabalin	Vigabatrin*
7	Zonisamide	Zonisamide	Tiagabine*
8	Vigabatrin	Tiagabine	Oxcarbazepine*
9	Lamotrigine	Lamotrigine	Lacosamide*
10	Gabapentin	Topiramate	Zonisamide*
11	Lacosamide	Oxcarbazepine	Topiramate*

4.6.3 Principal Findings

This systematic review, frequentist meta-analysis and network meta-analysis with supportive quantitative trade-off analysis provides the most comprehensive and explicit assessment of the short-term comparative efficacy and tolerability of AEDs licensed for the adjunctive management of chronic refractory partial epilepsy with or without secondary generalization. The frequentist meta-analysis demonstrated that all AEDs were superior to placebo in preventing seizures. However, there was a dearth of published, randomised clinical trials that directly compared one AED with another. Using a Bayesian hierarchical model to conduct a mixed-treatment network meta-analysis, levetiracetam, vigabatrin, sodium valproate and gabapentin emerged as agents with the best combination of short term efficacy and tolerability, with the caveat that vigabatrin is recognized as being associated with serious visual disturbance with chronic use.

4.6.4 Relation to Previous Studies

Five meta-analyses and a narrative review exist related to this question, which have produced conflicting results, partly due to differences in methodology and inclusion criteria. Marson *et al*. (89, 93) included 13 published and 15 unpublished randomised controlled trials, comparing six AEDs against placebo (gabapentin, lamotrigine, tiagabine, topiramate, vigabatrin, and zonisamide), concluding that there was a lack of conclusive evidence to determine a prescribing hierarchy accounting for differences in efficacy or tolerability. Shorvon *et al*. (143) subsequently included 36 published articles (including randomized controlled trials in full and abstract form), comparing eight AEDs against placebo, again reaching the same conclusion. Otoul *et al*. (144) later updated the meta-analysis conducted by Marson *et al*. by comparing the six AEDs indirectly against levetiracetam. No restrictions were placed regarding age of patients. Unlike the previous two reviews, the authors here found levetiracetam to demonstrate significant efficacy compared with gabapentin and lamotrigine. No significant differences regarding tolerability profile were noted between the seven AEDs evaluated.

In contrast, the present analysis, using a treatment network, did not demonstrate levetiracetam to possess a significantly more effective profile although it was one of four AEDs which showed a trend towards better efficacy and tolerability.

Beyenburg *et al*. (145) conducted a meta-analysis of modern AEDs to determine the true effect of each agent in reducing seizure frequency not attributable to other factors, principally placebo response. Although the search period was similar to that of the present study, the inclusion of both adults and children resulted in a larger number of eligible studies. The key finding of this paper was the estimation of a placebo-corrected difference of 21% (95% CI 19% to 24%; *P* <0.0001). This is akin to value calculated within the present study of 16%; the difference is likely to be the result of differences in the inclusion criteria.

Rheims *et al*. (146) more recently conducted a meta-analysis in adult patients investigating the different parameters which may determine response to treatment. The number of publications and AEDs included were again greater than the present study as a result of the inclusion criteria permitting AEDs currently under investigation, or not indicated for refractory focal epilepsy. The authors concluded that although responder rate increased over the years for the placebo arm, a parallel increase was also observed in the active arm. Further, the use of last observation carried forward (LOCF) data overestimated the responder rate and analysis of efficacy at the level of doses for each AED did not reveal statistically significant differences.

Lastly, Devinsky and Cramer commended the use of meta-analysis as the best available technique in the absence of comparative trial data.(147) However, this predated the availability of the more advanced network meta-analysis methodology. As such, the present analysis utilised a common placebo-corrected value, calculation of intention-to-treat dataset, and assessment of each AED at the most clinically relevant dose incorporating a Bayesian paradigm, an approach which is recommended for its clinical relevance,(148) with a

trade-off analysis to determine a treatment hierarchy. Prior overviews have supported the validity of network meta-analysis for indirect treatment comparisons provided certain conditions are met,(149) with emerging examples of its application in several disease areas.(150-154)

4.6.5 Strengths and Weaknesses of the Methodology

Using frequentist methodology, previous attempts to determine a comparative review were unsuccessful.(145) This was the first analysis within the epilepsy field to utilise Bayesian statistics. Relevant trials were identified via explicit systematic review and the analysis conformed to PRISMA recommendations.(25) The efficacy outcome selected within this analysis is universally accepted as an informative outcome measure concerning AED efficacy (CHMP/EWP/566/98 Rev. 2).(155) All studies included in the present analysis were fully published unlike the previously published meta-analyses (Marson et al.)(89) where over 50% (15/29) of included studies were unpublished; this stricter criteria was included to minimise the risk of heterogeneity as full trial methodology could be determine upon assessment for inclusion. Furthermore, the present analysis only included agents which are currently licensed in the UK for the adjunctive treatment of refractory epilepsy to strengthen the clinical utility of the findings generated. The criterion of single-dose extraction for analysis was supported by Rheims et al.(146) Lastly, a multiple treatment comparison design was utilised which allowed both direct and indirect comparisons via the construction of a treatment network unlike frequentist meta-analysis.(156)

Multiple treatment options exist in epilepsy but, in common with many other therapeutic areas, treatment comparisons may be missing and / or there may be preferences toward certain comparators (including placebo).(157) These were features of the current analysis where of the 65 possible comparisons between the 11 AEDs (and placebo), only 13 were conducted and all but three were placebo-controlled trials. A major reason for the lack of direct active comparator trials is that regulatory authorities such as the Committee for Medicinal Products for Human Use (on behalf of the EMA), do not require such

trials to be conducted as a condition of the marketing authorisation of a new drug. Note, this distinction is different to denying patients to any form as treatment as subjects included within the RCTs were on a background of AEDs; however the comparator was placebo rather than an active agent.

The recently published guideline on clinical investigation of medicinal products in the treatment of epileptic disorders continues to recommend that add-on studies (the addition of a new AED to existing therapies) should be of randomised, double-blind, placebo-controlled, parallel group study design,(155) although the guideline also recommends that as more AEDs are approved for the add-on indication, comparative trials may be considered and that an evaluation of efficacy of such agents may be conducted through meta-analysis. Despite the above recommendation, AEDs continue to be trialled and receive regulatory approval on the basis of placebo-controlled studies (with background treatment permitted), as highlighted by recent approval of eslicarbazepine and retigabine which continue to compound the situation. In the absence of prior active-comparator trial, guidance from NICE for example does not make recommendations about the selection or sequence of AED therapy with regard to specific drugs within the categories of older and newer AEDs.(90) An additional and crucial problem which arises from the absence of a prevailing treatment hierarchy in the presence of multiple therapeutic options is that the number of legitimate active comparator trials for any new AED becomes very large indeed, in itself providing an additional obstacle to comparator trials of new agents. Thus indirect comparisons, through network meta-analysis such as the one conducted here, might not only help develop rational treatment hierarchies based on existing therapies but also inform on the choice of high priority comparator agents for future trials of new AEDs.

Nevertheless, it is also important to take note of several important limitations. First, although individual trials only provide information over a short period of time (typically 8 to 16 weeks), this is the duration of follow-up required to meet regulatory criteria. The findings

should be extrapolated to longer term use with caution, particularly as some drugs, for example vigabatrin, are recognized as being associated with the development of adverse effects with long term use.

Second, the trials included in this analysis spanned 19 years during which time the available services and the management of epilepsy has changed. Importantly, the patient population recruited to more recent trials may have more refractory or 'drug resistant' epilepsy than patients recruited to earlier studies who were exposed to fewer previously licensed therapies. Also, it is possible that the aetiology of epilepsy has changed over the period 1990-2010; for example, epilepsy is an associated consequence of head traumas secondary to road traffic accidents, an issue which has been addressed by improved road safety. Statistical tests for heterogeneity (tau-squared and I-squared) considered that the patient populations were sufficiently similar to generate overall estimates, values noted as mild-moderate and moderate, respectively for the efficacy and tolerability endpoints. This point has been investigated in detail by Beyenburg *et al*.(145) who recently conducted a meta-analysis of modern AEDs, concluding that meta-regression showed no difference in effects between studies published before 2001 *vs.* 2001 and later for 50% reduction in seizure (Z = -0.50, *P* = 0.62) or seizure freedom (Z = -0.52, *P* = 0.60).The year 2001 was chosen to reflect the comparison of first *vs.* second generation AEDs. Beyenburg *et al*. also sought to explore whether or not factors such as number of past AEDs trialled or highest dose studied *vs.* all available doses had an influence on seizure-free outcome. However they concluded that insufficient data were available from the published manuscripts to enable such analyses. Nonetheless, the inclusion of multiple doses for each AED, specifically those at the higher end, may potentially skew the overall results with regards to increasing the incidence of premature discontinuation due to treatment-related adverse events without a corresponding increase in responder rate. Such heterogeneity would appear on Forest plot analyses as wider confidence / credible intervals and may reduce the validity of the results and limit generalisation of the findings. On this basis, despite a

reduction in overall patient numbers within the meta-analysis, it was decided to include data only for the dose closest to that prescribed in clinical practice for each AED.

Third, it is plausible that using the target of 50% reduction in seizure as the efficacy endpoint for a refractory cohort is unjustified and not reflective of clinical practice, for example, in patients who have failed to respond to 2-3 medicines the clinical community would suggest that a lower threshold of response be considered acceptable. Currently however, the 50% responder rate is an endpoint which is endorsed by regulatory authorities such as the EMA, which does not take into account the duration and / or severity of seizures. As such, any analyses would be required to conform to these criteria.

Fourth, although it was an intention from the outset to restrict the analysis to the adult population, a number of studies actually included a mixed population of adults and children (with one as low as ≥ 2 years) without sub-grouping their results. Nonetheless, as the majority of data relates to investigation in adults the conclusions drawn should remain relevant to the use of AEDs in the adult population only.

Fifth, this analysis only provides a non-specific estimate for withdrawal rates due to an intolerable adverse effect rather than firm conclusions on the severity of a specific adverse effect, such as ataxia, or the incidence of serious / rare adverse effects such as Steven-Johnson syndrome or suicidal tendency. Furthermore, although the inclusion criteria specified a double-blind design to promote robustness of the findings, it is appreciated that patients randomised to active therapy may become aware that they are receiving active treatment due the emergence of treatment-related adverse effects, such as sedation.

Sixth, the studies included may be limited in their ability to report on adverse effects as the studies were primarily designed and powered to address the effectiveness of the AED under investigation and had a maximum duration of 24 weeks.(158)

Seventh, in specifying such strict inclusion criteria to ensure higher quality trials were included in the quantitative analysis, it was necessary to exclude trials investigating AEDs used in routine clinical practice today, such as carbamazepine.

Finally, data within this analysis only relate to the use of AEDs for the treatment of patients with chronic refractory partial epilepsy and should not be extrapolated to patients with other forms of epilepsy, such as primary generalised. As for any medication, prescribing decisions should take into account specific contraindications and advice on prescribing in special groups such as women of childbearing age.

The present study included data from publications in English language only. The decision to not include data from publications in languages other than English (LOE) was made based on the conclusions reached by the Canadian Agency for Drug and Technologies in Health (CADTH) and the UK National Institute of Health Research (NIHR) (review of impact of language restrictions on systematic reviews). First, the CADTH performed a systematic review of meta-analyses of conventional medicines, comparing those that did and those that did not include publications in LOE concluding that they could find no evidence of a systematic bias from the use of language restrictions.(34) Second, a Health Technology Appraisal (on behalf of the NIHR) also found that language restrictions did not bias the results of systematic reviews of conventional medicines, even after sensitivity analyses were conducted, noting that the results do not appear to be influenced by statistical heterogeneity or publication bias.(159)

Finally, as carbamazepine is widely recommended and accepted as being the first-line option for patients with partial-onset seizures,(160) we believe that the non-inclusion of this particular AED within this analysis should not limit its findings.

4.6.6 Implications for Clinical Practice

This review is focused on the most common form of epilepsy seen in clinical practice, partial onset seizure with or without secondary generalisation in the refractory setting. The stimulus for this stems from the lack of a guideline which provides explicit detail on how each agent licensed for this indication should be prescribed relative to one another.

Although it is widely accepted that the best model of providing answers with the highest clinical relevance would be from a large-scale multiple comparison head-to-head trial, the costs associated with such a model render the chances of this transpiring to be very small. The present study attempts to provide a response to one of the limitations of current trial methodology as specified within the discussion of the recent publication by Beyenburg *et al*. (145) – 'given the absence of double-blind, placebo-controlled comparative AED trials, we could not compare the efficacy among individual AEDs' – by generating a treatment network utilising a Bayesian framework supplemented by a quantitative trade-off analysis using pre-specified efficacy and tolerability parameters as well as acknowledging the current acquisition costs of each agent. In order to generate a robust and simplified network from which to permit translation into a treatment hierarchy, a single dose per AED was extracted for the analysis. This strategy follows the results of Rheims *et al*. (146) where an analysis of dose selection resulted in mostly non-significant difference for the primary endpoint (50% reduction in seizure frequency). In light of these data, such analyses were not repeated and instead data reporting for the dose investigated which most closely reflects that in current clinical practice were extracted for each AED.

Until conduct of an independent large-scale RCT, herein the chapter is reported a methodologically and statistically rigorous analysis of AEDs currently available in the UK to assist clinical decision-making. An overview of the strengths and weaknesses of the methodology used is detailed above in section 4.6.5.

Lastly, an important implication of the findings is that although it is clear that the modern AEDs are more effective than placebo when used as adjunctive therapy for refractory focal epilepsy, they are of limited efficacy (in comparison with older agents) and are correlated with an increase in treatment-related intolerable adverse effects as well as cost. The current strategies for developing, investigating and licensing of AEDs thus require re-evaluation.

Interestingly, this area of work continues to be of interest as evident by an NMA undergoing editorial review by the Annals of Medicine, July 2017 (SANN-2016-0302). In the submitted analysis, the authors aim to compare the relative efficacy and tolerability of the second and third generation AEDs for refractory epilepsy using NMA methodology. The 50% responder rate was again selected as the efficacy outcome whilst the tolerability outcome was more focussed on the exploring the incidence of dizziness and somnolence rather than a more general withdrawal as used in the present analysis. Presentation of outcome was again in the format of odds ratio (OR) and their 95% credible interval (CrI) as obtained from NMA, however ranking was determined via an analysis of surface under the cumulative ranking curve (SUCRA), which is an alternative strategy to a plot of number needed-to-treat (NNT) vs. number needed-to-harm (NNH). The outcomes were broadly similar to the current study with topiramate suggested as being significantly more effective than placebo, eslicarbazepine acetate, perampanel, pregabalin, zonisamide, gabapentin and lamotrigine with respect to 50% RR (all OR > 1). With regards to tolerability, patients who were managed by eslicarbazepine acetate, perampanel, oxcarbazepine, topiramate and pregabalin were more likely to suffer from dizziness compared to those who receive placebo (all OR > 1). Levetiracetam appeared to possess the best overall balance of efficacy and tolerability (SUCRA = 0.769, 0.743, 0.604 and 0.659), which concurs with the present analysis (see Figure 25, which plots levetiracetam in the lowest right hand corner indicating the best trade-off as lowest NNT and highest NNH profile).

Although the authors of the recent [currently unpublished] study utilised a slightly different methodology for determining the trade-off between efficacy and tolerability i.e. SUCRA vs. a plot of NNT / NNH, the results are in keeping with the findings of within this chapter. It is also interesting that independent interest has been shown in a variant trade-off assessment within the same field and specific area as presented here. The results of the present project therefore remain relevant with the recent NMA adding little to this field than previously established here.

4.7 Summary

This systematic review and Bayesian network meta-analysis highlights the paucity of long-term prospective randomised active comparator trials of AEDs currently licensed for refractory epilepsy. In comparison with the five previously published meta-analyses reporting on this topic and building on the recommendations issued by NICE, the use of indirect treatment comparisons indicated that levetiracetam, vigabatrin, gabapentin, and sodium valproate demonstrated the best combination of short term efficacy and tolerability, whereas oxcarbazepine, while equally effective, was the least well tolerated in the short term.

As the use of carbamazepine is widely recommended and accepted as being an effective first line option for patients with partial onset focal seizures, despite there being no trials investigating this agent within the present analysis, clinical practice has dictated this agent to demonstrate a good balance of efficacy and tolerability in both the short and long term.

With the exception of vigabatrin which is associated with visual field defects in long term use, in the absence of evidence definitively indicating the clinical superiority or tolerability of one AED over another, we suggest that other factors such as acquisition cost, dosing regimen, licensing indications and contraindications be the differential in selecting among these AEDs.

The logical next step for future trials would be the commissioning of multiple active comparisons, similar to the SANAD study in primary

generalized epilepsy and partial epilepsy.(86, 87) Until regulators mandate greater use of such trials, network meta-analysis incorporating mixed treatment models and the hierarchical trade-off analysis may provide the only available source of useful information on comparative efficacy and safety of treatments for epilepsy and other common diseases.

An independent review exploring the characteristics and methodological quality of network / indirect meta-analyses conducted over the last 10 years has recently been published.(161) As part of this analysis the quality of the above publication has been graded as moderate – high in accordance with the AMSTAR (a measurement tool to assess systematic reviews) and ISPOR (international society of pharmacoeconomics and outcomes research) criteria.

This work has been published in the British Journal of Pharmacology 2013; 76(5): 649-667 (doi: 10.1111/bcp.12083) and meets the criteria for inclusion on the NIHR Centre for Reviews and Dissemination (CRD) Database of Abstracts of Reviews and Effects (DARE).

4.8 Implications of Research

This publication utilised a novel trade-off analysis to comparatively determine the key efficacy and safety / tolerability parameters of treatments for refractory epilepsy in the absence of a large scale all-encompassing randomised controlled clinical trial. The findings from this area of research were published and communicated locally via the area prescribing committee. The results have been read with interest and the NMA model has served as a template / reference point within the region for the use of new AEDs that have come to the UK market since, including eslicarbazepine, retigabine and perampanel (see Figure 26), all of which appear to possess a similar (and not superior) efficacy profile to those previously available (as compared with Figure 17).

The introduction of newer agents, as demonstrated with eslicarbazepine as a variant of oxcarbazepine, within North Central London are therefore more rigorously assessed with a conservative

uptake compared with adjacent regions as a result of this research project (see Figure 27). A locally developed prescribing hierarchy lists (following the failure of first-line gabapentin or valproic acid) the preferred second-line choices being levetiracetam > topiramate > lacosamide > zonisamide with consideration to the balance of NNT vs. NNH and cost.

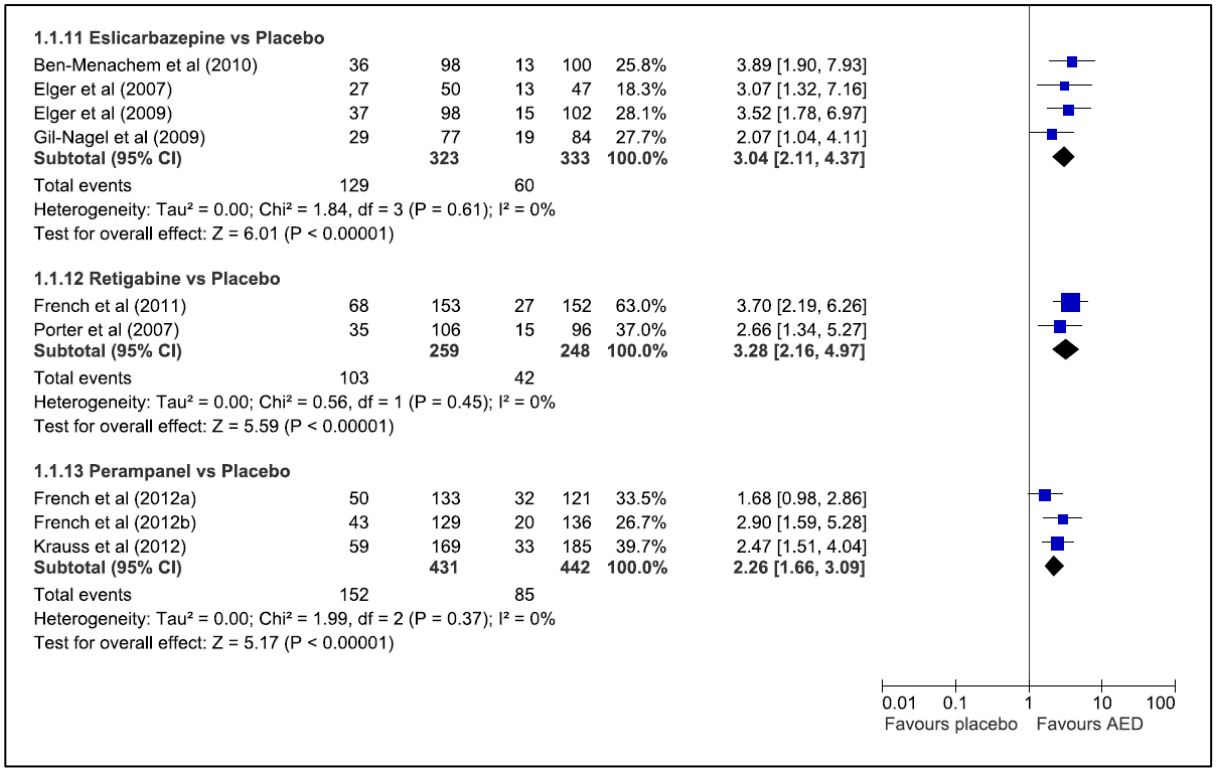


Figure 26: Expanded analysis forest plot (RevMan v5.0) of the odds ratios for efficacy (50% responder rate) of randomized controlled trials comparing an AED *vs.* placebo as add-on treatment for refractory epilepsy, respectively.

The black squares represent the odds ratio for individual studies of AED *vs.* placebo and the horizontal line represents the 95% confidence interval of the odds ratio. The black diamond represents the random-effects pooled odds ratio for studies reporting on the same AED where its width represents the 95% confidence intervals. Estimates to the right of the vertical line (i.e. odds ratio >1) are indicative of a statistically significant increase in efficacy, relative to placebo, in patients randomized to the active intervention.

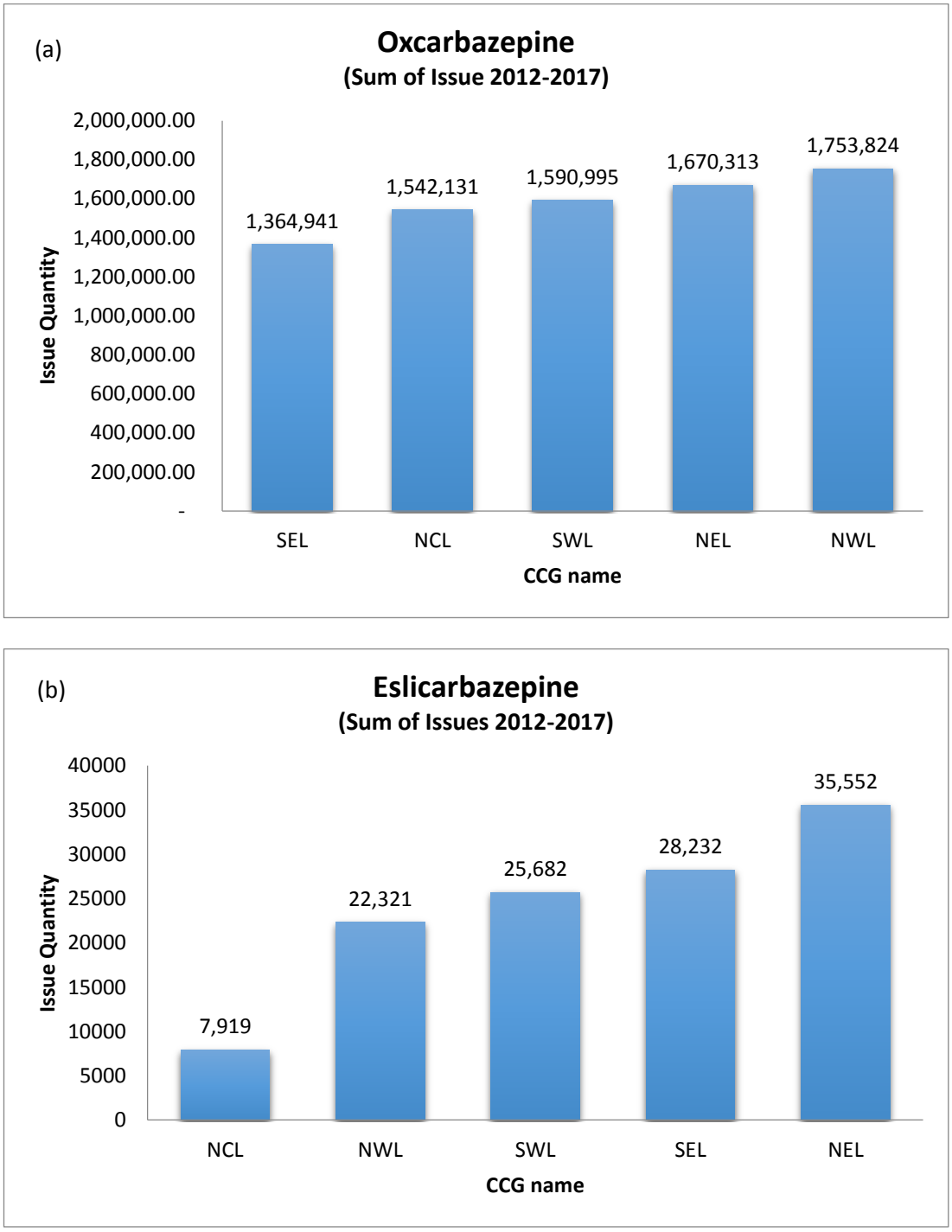


Figure 27: Sum of issue quantity (June 2012 – May 2017) for (a) oxcarbazepine, and its newer equivalent (b) eslicarbazepine, within general practice across the London health economy

Based on the NMA findings in Figure 26, where the efficacy (50% responder rate) for eslicarbazepine is lower than that for oxcarbazepine (Figure 17), the NCL area prescribing committee could not justify the addition of eslicarbazepine to the local formulary. The total issue quantity for eslicarbazepine within NCL compared with the other four London CCGs is therefore considerably lower. Source: **https://openprescribing.net**

[BLANK PAGE]

5.0 Chapter 5: Research Project Two – Angiotensin-II Receptor Antagonists for the Management of Hypertension and Heart Failure

5.1 Introduction

The second research project introduces the concept of a cost-utility analysis as the trade-off assessment was expanded from efficacy and safety / tolerability to efficacy and cost. A systematic review and frequentist meta-analysis of all published studies will be performed, comparing losartan and candesartan as the two leading angiotensin II receptor antagonists for the management of hypertension and heart failure. Data will be abstracted from RCTs of head-to-head design to obtain estimates on their relative efficacy and safety with a view to determine the trade-off with cost-effectiveness. The use of a frequentist rather than Bayesian meta-analysis will be suitable for this project as only trials investigating both agents in a direct comparison will be identified. Further, the cost-utility method will be suitable as both agents are from the same pharmacological class and therefore regarded as clinically interchangeable.

The goal for this project is to disseminate the outcomes via NHS channels to influence prescribing across the UK, including guides for clinicians and patients.

5.2 Literature Review

5.2.1 Cardiovascular Disease

Cardiovascular disease (CVD) is the leading cause of morbidity and mortality worldwide (162), despite advances in diagnosis and management. Hypertension is a major risk factor for CVD, and a risk factor for the development of heart failure. The goal of hypertension therapy is to reduce blood pressure (BP) to less than 140/90 mmHg in elderly patients and to less than 130/85 mmHg in the young or middle-aged and in patients with diabetes mellitus irrespective of age (163). Patients with symptomatic chronic heart failure (CHF) have reduced cardiac output, characteristic symptoms of dyspnoea, orthopnoea and decreased exercise capacity, and have a high risk of death and hospitalisation (164).

5.2.2 Current Management

Although angiotensin-converting enzyme (ACE) inhibitors are a well-established class of treatment, some patients are unable to tolerate them. In recent years, angiotensin II receptor antagonists (AIIRAs) have emerged as an alternative option for targeting and inhibiting the renin-angiotensin-aldosterone system (RAAS) by selectively blocking the AT1 subtype (165). AIIRAs exert a similar antihypertensive effect to ACE inhibitors (166, 167); however, their specificity avoids major ACE inhibitor-related adverse effects, such as cough and angioedema, which are believed to result from non-specific interference of bradykinin metabolism (168). Compared with losartan (Cozaar; MSD, Whitehouse Station, NJ, USA), the first AIIRA to receive a marketing authorisation for the management of hypertension, candesartan (Amias; Takeda, New York, NY, USA), the current market leader, has a slower dissociation rate from the AT1 receptor (169), potentially providing it with a longer-acting antihypertensive effect.

5.3 Rationale for this Research Project

The UK National Health Service (NHS) in 2010 spent in excess of £250 million per annum on AIIRAs for the treatment of hypertension and heart failure; with candesartan cilexetil (Amias®; Takeda) dominating the market. Until recently both agents were comparably priced, however, as losartan potassium was the first AIIRA to receive a Marketing Authorisation, generic preparations of this agent are emerging at a fraction of the cost to its branded equivalent, Cozaar® (MSD). This research project explores the potential health benefits and compromises that a trade-off for a cheaper generic medicine would bring, through an exploration of its cost-effectiveness compared with a more expensive ‘on patent’ alternative.

5.4 Collaboration

All work related to the systematic review and frequentist meta-analysis within this chapter of the thesis was undertaken by me. Systematic review support was provided by Dr Grosso in the form of a second independent check against the proforma. The cost-utility model was

built and executed by Dr Scott with clinical information obtained from the meta-analysis. Subsequent analyses were interpreted and circulated by me with support from Dr Grosso.

5.5 Protocol

5.5.1 Systematic Review

Relevant randomised trials were searched for within the Cochrane Central Register of Controlled Trials (Cochrane Library 2009, issue 2), which contains the Hypertension and Heart Group's specialist register, Medline (1950–March 2010), and Embase (1980–February 2010). The search terms and limits are provided below (see section 5.5.4).

In addition to the database search strategy, the reference lists of identified manuscripts were manually hand-searched to identify additional relevant studies. To formulate and ensure optimal reporting of the systematic review and meta-analysis, the established PRISMA statement was followed (170). A summary of the studies identified, screened, assessed for eligibility and included for analysis is summarised in Figure 28 and Figure 29 for the hypertension and heart failure indications, respectively.

5.5.1.1 Inclusion Criteria

Hypertension

Trials were included if they were of randomised, double-blind, active-controlled design (investigating candesartan and losartan), provided they recruited adult patients (> 18 years), were of parallel or cross-over design, and the treatment period was of at least 4 weeks duration.

Heart Failure

Trials were included if they were of randomised, double-blind, active-controlled design (investigating candesartan and losartan), provided they recruited adult patients (> 18 years) with a left ventricular ejection fraction (LVEF) not exceeding 40%, were of

parallel or cross-over design, and the treatment period was of at least 4 weeks duration.

5.5.1.2 Exclusion Criteria

Studies were excluded if they were open-label, observational, or not fully published (e.g. only presented at conference proceedings or solely available in abstract form). Non-English language publications were also excluded (in accordance with the findings described in section 4.6.5) as were trials which permitted the use of other therapies which may have confounded the clarity of the outcomes of the drug being assessed (e.g. calcium channel blockers). Trials which used a conditional design where patients were allocated treatment only if they showed a pre-determined response to treatment during a baseline period before randomisation were also excluded.

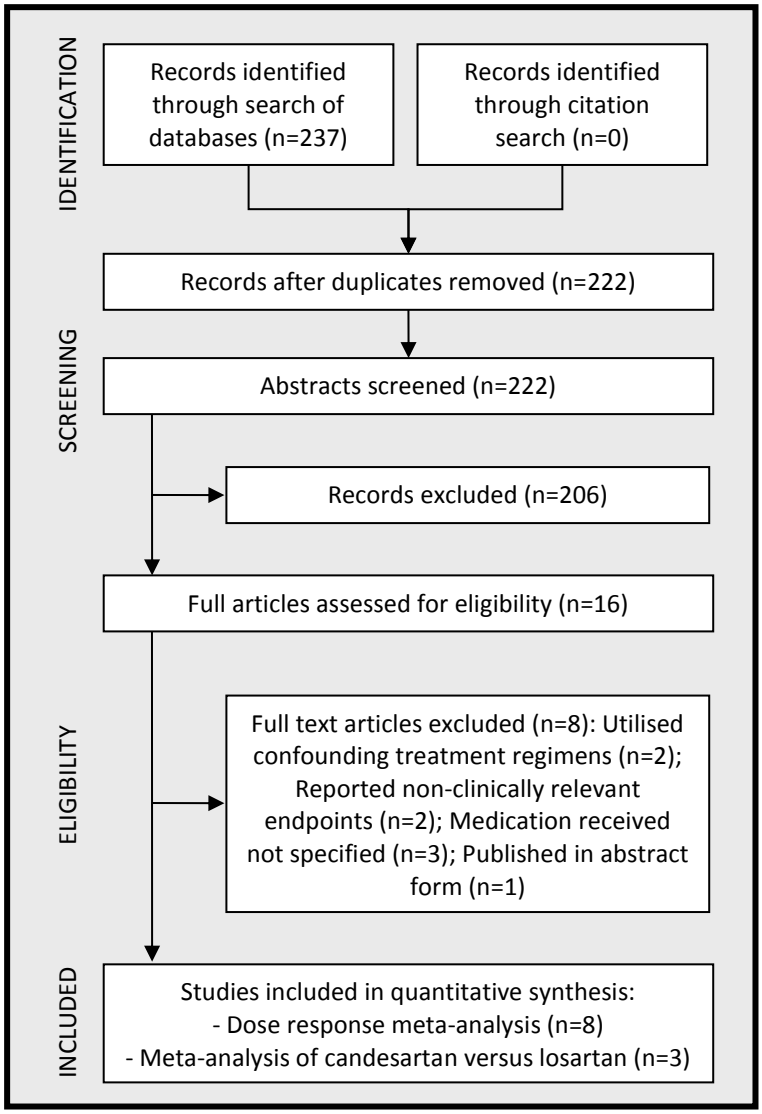


Figure 28: The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of studies identified from the systematic review for inclusion in the meta-analysis for hypertension

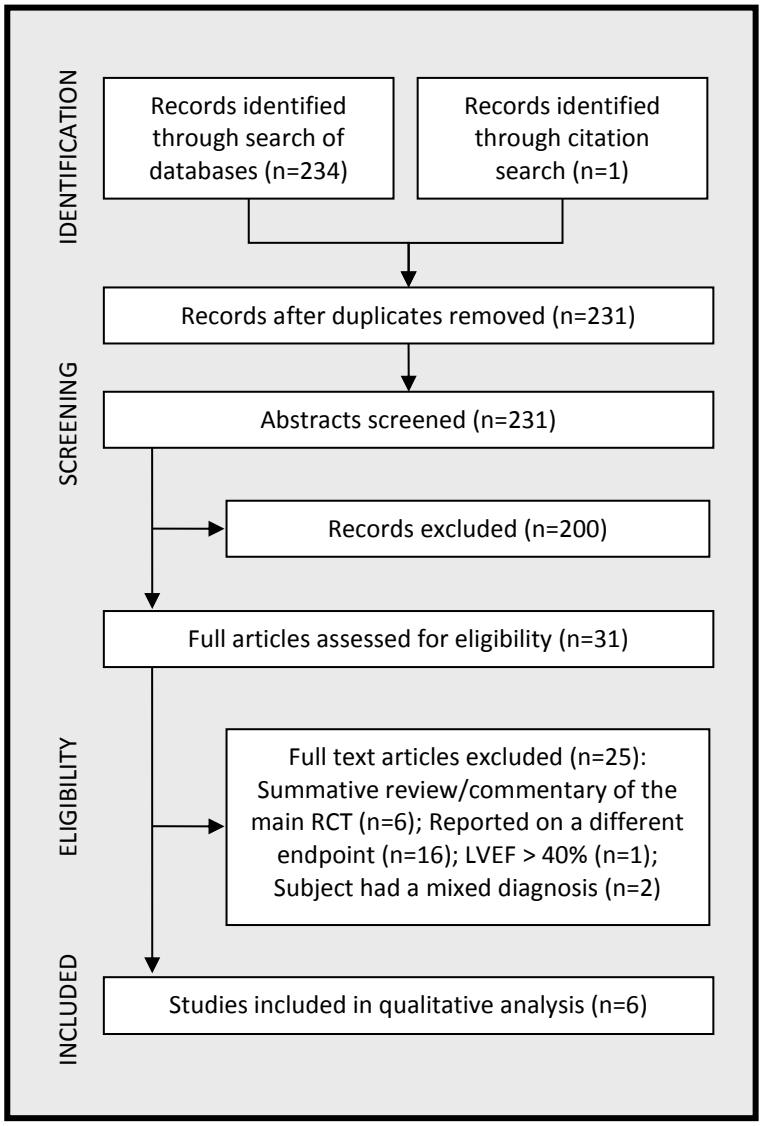


Figure 29: The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of studies identified from the systematic review for inclusion in the meta-analysis for heart failure

5.5.2 Endpoints

The primary efficacy endpoints for extraction from the hypertension trials were the mean change from baseline in trough (24 hours post-dose) diastolic blood pressure (DBP) and systolic blood pressure (SBP). The primary efficacy endpoint for extraction from the heart failure trials was a composite of cardiovascular death and hospitalisation admission for management of heart failure. As safety / tolerability profiles for the two agents overlap due to them being from the same class, this endpoint was not assessed.

5.5.3 Quality Assessment and Data Collection

Data from systematic reviews or previous meta-analyses were not used to enable the collection of data from original sources; however, any such publications identified served as a comparator to ensure that all relevant studies had been included within this review. Secondary searches were conducted from the reference lists of manuscripts identified. To minimise bias, references and abstracts retrieved by the search were reviewed by me and validated by Dr Grosso. The same process applied to the selection of potentially relevant publications against the pre-specified inclusion and exclusion criteria. Important clinical and methodological study characteristics were extracted onto a standard form, checked and recorded to ensure consistency. This included:

1) Characteristics of trial participants (including age, gender, ethnicity, and diagnosis)
2) Study design and regime of intervention
3) Details of intervention (including type, dose, and duration)
4) Details of outcome measure (baseline, peak and trough for SBP and DBP).

Any discrepancies or lack of agreement between the two reviewers were referred to a third independent investigator (ADH) for arbitration. An assessment of risk of bias (using established criteria) (27) was also undertaken. All analyses were based on *intention-to-treat* data. For any trials that reported data using a *per-protocol* analysis, *intention-to-treat* values were calculated.

5.5.4 Search Terms

See Table 10 and Table 11 for a list of the search terms and limits applied when conducting the database search.

Table 10: Search strategy used to determine eligible published trials for the hypertension and heart failure analysis

	Search term
01.	Losartan [MeSH]
02.	Candesartan$.tw
03.	Candesartan cilexetil [MeSH]
04.	Hypertension [MeSH]
05.	Heart failure [MeSH]
06.	1 AND (2 OR 3) AND 4
07.	1 AND (2 OR 3) AND 5
08.	6 OR 7

Table 11: Limits applied to each database as part of the systematic review for the hypertension and heart failure analysis

Database	Limits applied
PubMed	Humans; Randomised controlled trial; English; all adult: 19+ years
Cochrane	Nil
Embase	Full text; Human; English language ; Article OR Erratum; Adult <18 to 64 years> OR aged <65 years+

5.5.5 Statistical Analysis – Frequentist Meta-Analysis

A random-effects 'frequentist' meta-analysis (38, 171) was conducted, to compare outcomes, reported as the weighted mean difference; therefore estimates meta-analysed over multiple trials are considered as average treatment effects. The choice of random-effects is in keeping with previous analyses where the protocol and patient population for each study included vary slightly.

To evaluate heterogeneity of the effect estimates the Cochran Q (Chi-squared) and Higgins I-squared statistics were calculated (48). Refer to section 3.5.7 for further details on these.

To generate the Forest plots, Funnel plots and l'Abbé plots, StatsDirect® v.2.7.7 (Altrincham, Cheshire, UK) was used.

5.5.6 Economic Analysis – Cost-Utility Analysis (CUA)

[This section of the research project was performed by MS]

A 10-year CUA of adopting losartan or candesartan using a Markov state transition model (see Figure 30) was developed using Microsoft Excel (Microsoft Corporation, 2007) by Dr Scott in consultation with Dr Grosso and me. Structurally, the model assumes that all patients can be in one unique health state in a given cycle [one of: 'Well', 'Coronary Heart Disease (CHD)', 'Stroke' and 'Death'] with the entire cohort initially in the 'Well' state. A 10-year time horizon and annual cycle length was used for comparison of the costs and health outcomes for patients with essential hypertension, from the viewpoint of the UK National Health Service (NHS). Development of primary disease and subsequent quality-adjusted survival was considered in the model. Refer to section 10.3 (Appendix III) for further detail on the CUA model.

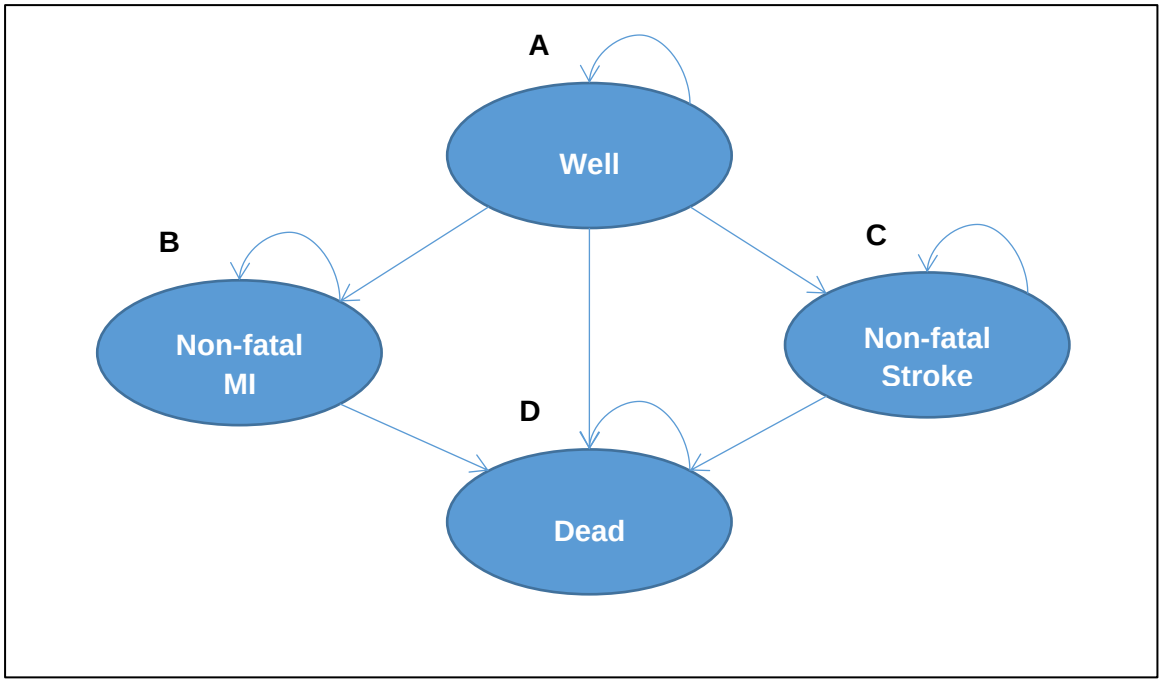


Figure 30: Markov State Transition diagram used for the cost-utility analysis.

The patient cohort all start in state A ('Well') and can transition annually to the coronary heart disease (CHD) and cerebrovascular disease states (B and C) denoted by the arrows, or they can survive or die from either myocardial infarction (MI) or stroke events from states B and C, respectively, or die from other causes. The time horizon of the model is 10 years

5.6 Findings

5.6.1 Hypertension

Overall, eight studies fulfilled the inclusion criteria for hypertension representing 9 comparisons and 3,619 patients (172-179). A summary of the characteristics of the studies included within this review is given in Table 12. Of the nine comparisons, two compared low-dose losartan with low-dose candesartan (50mg versus 8mg); two compared low-dose losartan with mid-dose candesartan (50mg versus 16mg); two compared high-dose losartan with mid-dose candesartan (100mg versus 16mg); and three compared high-dose losartan with high-dose candesartan (100mg versus 32mg). The primary efficacy endpoint data were extracted from these studies and pooled in a meta-analysis *(StatsDirect® 2.7.7; Altrincham, Cheshire, UK)* to estimate the weighted average reduction in DBP and SBP from baseline in the two treatment groups (see Figure 31 and Figure 32).

These analyses estimate a between-treatment difference of -1.89 mmHg (95% CI -2.29 to -1.48) for trough DBP and -2.96 mmHg (95% CI -3.60 to -2.32) for trough SBP in favour of candesartan. Overall, the nine comparisons generated an I-squared statistic of 32.6% for the trough DBP and 52.4% for trough SBP indicating that although the results are statistically significant, there is a mild-to-moderate degree of heterogeneity between the individual studies when combined.

For the purpose of the CUA, the meta-analysis was re-performed using only the three later studies which investigated both AIIRAs at their respective maximal licensed doses to derive a point estimate based on data reporting on comparative doses (see Figure 33 and Figure 34). This produced a between-treatment difference of -1.96 mmHg (95% CI -2.40 to -1.51) for trough DBP and -3.00 mmHg (95% CI -3.79 to -2.22) for trough SBP in favour of candesartan. As trough SBP is a more robust predictor of stroke and adverse cardiovascular outcome [Framingham model] this value was used the analysis.

Further detail of data extracted from the eligible studies to populate the meta-analysis are provided in Table 16.

Table 12: Characteristics of studies included within the meta-analysis for hypertension

Study	Study design (weeks of treatment)	Daily treatment dose (number of subjects)	Subjects			
			Diagnosis	Sex (Male:Female)	Mean age (range)	Ethnicity (Non-black:Black)
Hypertension studies						
Andersson OK and Neldam S (1998)(172)	Parallel, non-forced titration (8)	Losartan 50mg (83), Candesartan 8mg (82), Candesartan 16mg (84), placebo (85)	Mild-moderate	188:146	60 (20-80)	334:0
Lacourcière Y et al (1999)(177)	Parallel, Forced titration (8)	Losartan 100mg (115), Candesartan 16mg (116), placebo (37)	Mild-moderate	165:102	55 (20-80)	261:6
Gradman AH et al (1999)(175)	Parallel, non-forced titration (8)	Losartan 50/100mg (170), Candesartan 16/32mg (162)	Moderate	191:141	54 (18-80)	291:41
Manolis AJ et al (2000)(178)	Parallel, non-forced titration (12)	Losartan 50/100mg (461), Candesartan 8/16mg (458), Losartan+hydrochlorthiazide 50+12.5mg (232)	Mild-moderate	608:553	51 (20-80)	635:526
Bakris G et al (2001)(174)	Parallel, Forced titration (8)	Losartan 100mg (332), Candesartan 32mg (322)	Moderate	380:274	54 (18-80)	541:113
Vidt DG et al (2001)(176)	Parallel, Forced titration (8)	Losartan 100mg (304), Candesartan 32mg (307)	Moderate	358:253	55 (18-80)	490:121
Willemsen JM et al (2004)(179)	Cross-over, non-forced (4)	Losartan 50mg, Candesartan 16mg, placebo (total n=13)	Mild-moderate	8:5	52 (39-58)	NR
Baguet JP et al (2006)(173)	Parallel, non-forced titration (6)	Losartan 50mg (89), Candesartan 8mg (87), placebo (80)	Mild-moderate	153:103	54 (18-75)	NR

Table 13: Characteristics of studies included within the meta-analysis for heart failure

Study	Study design (weeks of treatment)	Daily treatment dose (number of subjects)	Subjects			
			Diagnosis	Sex (Male:Female)	Mean age (range)	Ethnicity (Non-black:Black)
Heart Failure studies						
Candesartan						
Granger CB et al (2003)(180)	Parallel (146)	Candesartan 32mg (1,013), placebo (1,015)	LVEF ≤ 40%; NYHA II-IV	1382:646	67 (NR)	1955:73
McMurray JJV et al (2003)(181)	Parallel (178)	Candesartan 32mg (1,276), placebo (1,272)	LVEF ≤ 40%; NYHA II-IV	2006:542	64 (NR)	2421:127
Granger CB et al (2000)(182)	Parallel (12)	Candesartan 16mg (179), placebo (91)	LVEF ≤ 40%; NYHA II-IV	186:84	65 (58-74)	NR
Losartan						
Konstam MA et al (2009)(183)	Parallel	Losartan 50mg (1,927), Losartan 150mg (1,913)	LVEF ≤ 40%; NYHA II-IV	2691:1149	66 (56-73)	NR
Pitt B et al (2000)(184)	Parallel	Losartan 50mg (1,578), Captopril 150mg (1,574)	LVEF ≤ 40%; NYHA II-IV	2185:966	72	NR
Pitt B et al (1997)(185)	Parallel	Losartan 50mg (352), Captopril 150mg (370)	LVEF ≤ 40%; NYHA II-IV	482:240	74	688:34

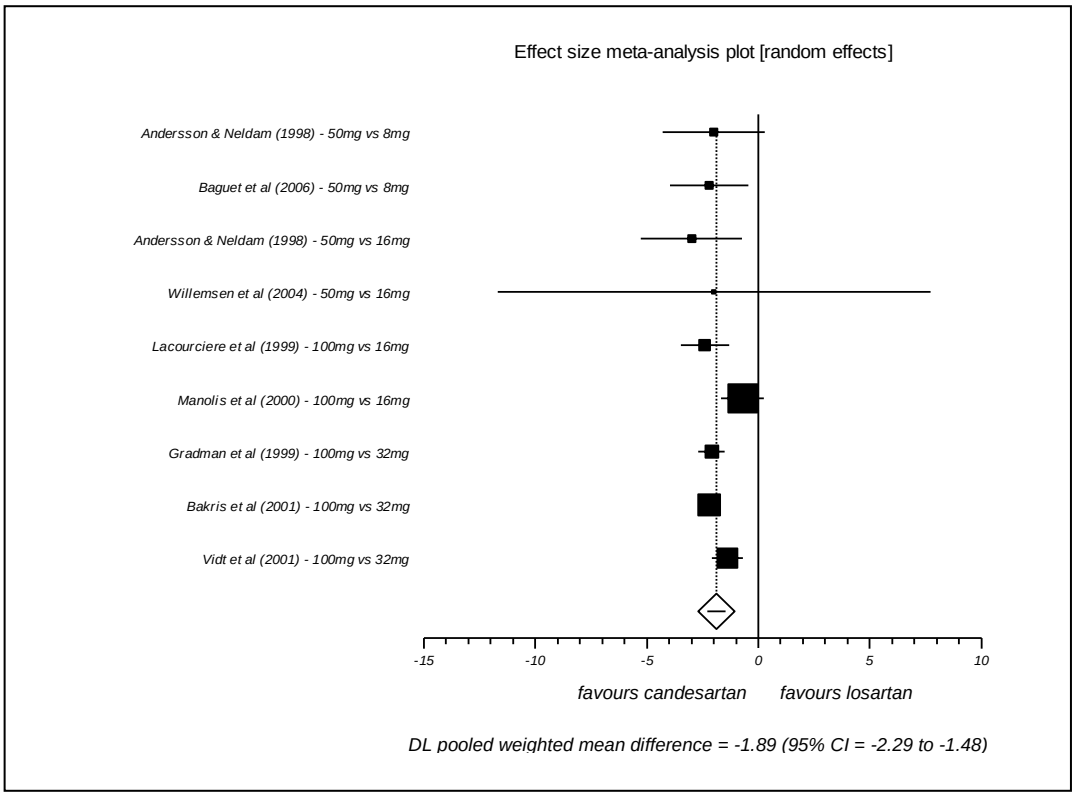


Figure 31: Forest plot of the weighted mean difference (WMD) for absolute difference in trough *diastolic* blood pressure from baseline to end of study period (4 to 12 weeks) of randomised controlled trials comparing losartan with candesartan as monotherapy for the treatment of mild-to-moderate hypertension.

The black squares represent the WMD for individual studies and the horizontal line represents the associated 95% confidence interval. The diamond and horizontal line within it represents the random-effects pooled WMD and its corresponding 95% confidence interval. Estimates to the left of the vertical line (i.e. WMD < 0) are indicative of a significant difference in trough *diastolic* blood pressure in favour of candesartan. Statistical significance is inferred where the confidence interval does not cross the vertical line of unity

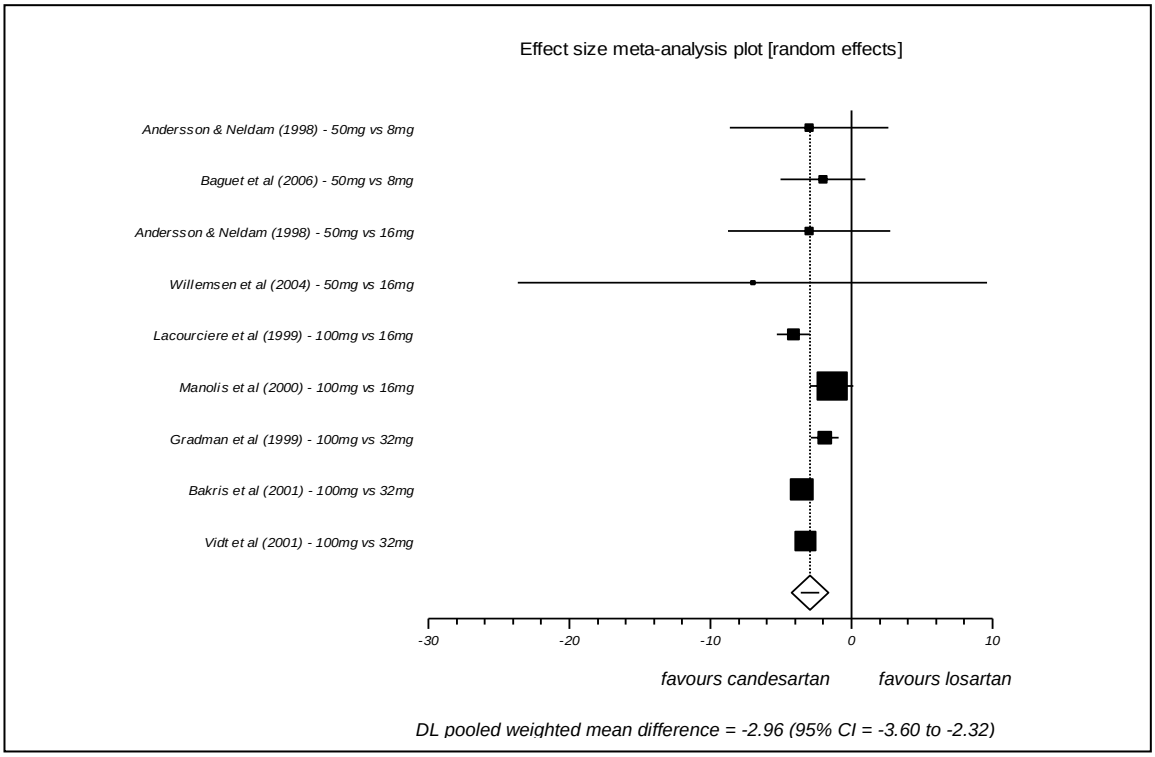


Figure 32: Forest plot of the weighted mean difference (WMD) for absolute difference in trough *systolic* blood pressure from baseline to end of study period (4 to 12 weeks) of randomised controlled trials comparing losartan with candesartan as monotherapy for the treatment of mild-to-moderate hypertension.

The black squares represent the WMD for individual studies and the horizontal line represents the 95% confidence interval of the WMD. The diamond and horizontal line within it represents the random-effects pooled WMD and its corresponding 95% confidence interval. Estimates to the left of the vertical line (i.e. WMD < 0) are indicative of a significant difference in trough *systolic* blood pressure in favour of candesartan. Statistical significance is inferred where the confidence interval does not cross the vertical line of unity

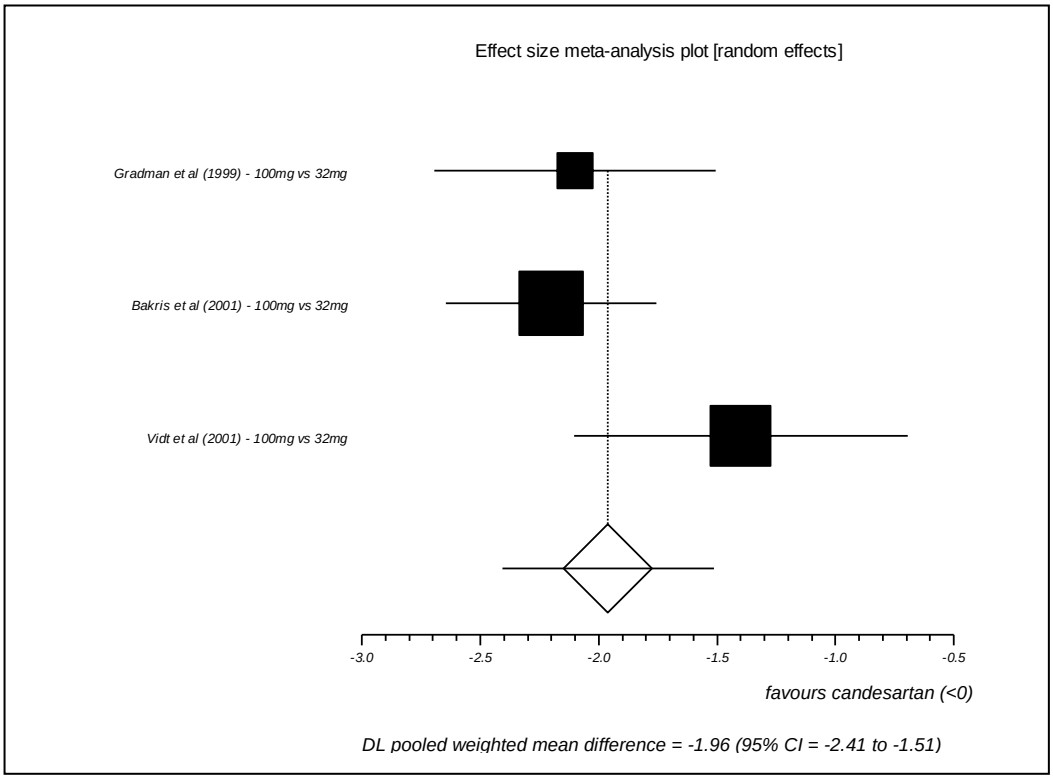


Figure 33: Forest plot of the weighted mean difference (WMD) for absolute difference in trough *diastolic* blood pressure from baseline to end of study period (4 to 12 weeks) of randomised controlled trials comparing losartan with candesartan *at their maximum licensed doses* as monotherapy for the treatment of mild-to-moderate hypertension.

The black squares represent the WMD for individual studies and the horizontal line represents the 95% confidence interval of the WMD. The diamond and horizontal line within it represents the random-effects pooled WMD and its corresponding 95% confidence interval. Estimates to the left of the vertical line (i.e. WMD < 0) are indicative of a significant difference in trough diastolic blood pressure in favour of candesartan. Statistical significance is inferred where the confidence interval does not cross the vertical line of unity

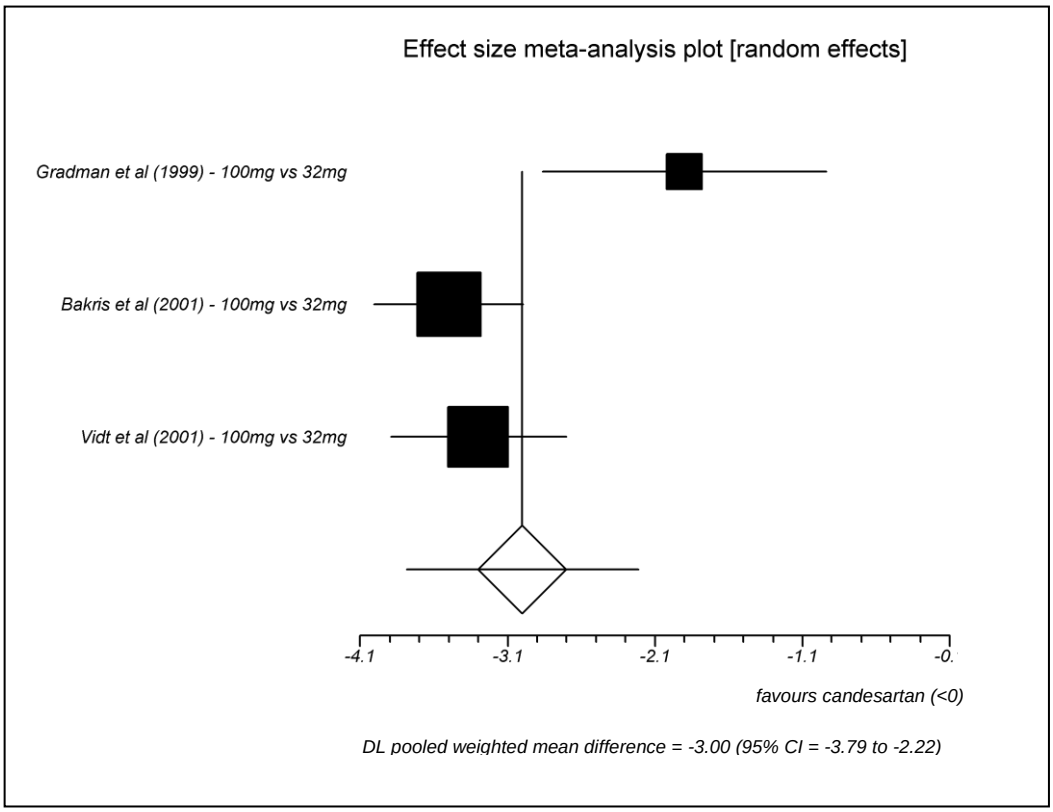


Figure 34: Forest plot of the weighted mean difference (WMD) for absolute difference in trough *systolic* blood pressure from baseline to end of study period (4 to 12 weeks) of randomised controlled trials comparing losartan with candesartan *at their maximum licensed doses* as monotherapy for the treatment of mild-to-moderate hypertension.

The black squares represent the WMD for individual studies and the horizontal line represents the 95% confidence interval of the WMD. The diamond and horizontal line within it represents the random-effects pooled WMD and its corresponding 95% confidence interval. Estimates to the left of the vertical line (i.e. WMD < 0) are indicative of a significant difference in trough *systolic* blood pressure in favour of candesartan. Statistical significance is inferred where the confidence interval does not cross the vertical line of unity

5.6.2 Base-Case Cost-Utility Model

The base-case Markov model used a generic losartan 100mg retail price of £6.47 (28-day pack price) compared with the list price of £16.13 for candesartan 32mg (28-day pack price) at a moderate baseline risk (SBP 165mmHg) in a cohort of men and women aged 65 years (see Table 14).

The difference in mean trough SBP between candesartan and losartan was obtained from the meta-analysis (-3.00mmHg). The estimated ICERs for male and female patients with 'moderate' risk were £44,930 and £53,804, respectively, demonstrating the cost-effectiveness of losartan relative to candesartan at current generic acquisition costs over a 10-year horizon.

5.6.3 Heart Failure

Overall, six studies were identified for heart failure; three studies reporting on losartan and three studies reporting on candesartan. No published head-to-head studies were found (see Table 13). In the absence of comparative trials directly comparing the efficacy of candesartan versus losartan in CHF, a qualitative analysis of the key studies was undertaken as an alternative to a quantitative meta-analysis (see section 5.7.2). As the only common comparator was placebo and there are no head-to-head studies comparing the two AIIRAs, a NMA to generate indirect comparative estimates was not considered appropriate.

Table 14: Cost and utility parameters used in the base-case utility model comparing candesartan and losartan for the treatment of hypertension

Parameter	Base-case value	Source
Annual drug cost of candesartan (32mg)	£193.56	BNF (2009)**(186)**
Annual drug cost of generic losartan (100mg)	£77.64	TEVA Pharmaceuticals **(187)**
Projected annual drug cost of generic losartan (100mg)	£10.56	Expert Opinion
Annual cost of stroke survivor [cost in first year]	£2,163 (8,046)*	NICE Guideline CG034**(188)**
Annual cost of MI survivor [cost in first year]	£500 (4,448)*	NICE Guideline CG034**(188)**
Utility weight of stroke survivor	0.63	NICE Guideline CG034**(188)**
Utility weight of MI survivor [first year]	0.88 (0.76)	NICE Guideline CG034**(188)**
Treatment effect difference (incremental reduction in SBP)	-3.00 mmHg	Meta-Analysis

*cost inflated to 2009 base case; (MI) myocardial infarction; (SBP) systolic blood pressure

Table 15: Variation of ICER with baseline risk when comparing candesartan and losartan for the treatment of hypertension

Baseline risk/SBP	ICER (£/QALY)	
	Generic losartan (£77.64 per annum)	Projected generic price (£10.56 per annum)
Males		
Mild (140 mmHg)	£52,644	£87,946
Moderate (165 mmHg)	£44,930	£74,901
High (180 mmHg)	£41,469	£69,076
Females		
Mild (140 mmHg)	£85,244	£142,449
Moderate (165 mmHg)	£53,804	£91,368
High (180 mmHg)	£41,591	£71,430

Table 16: Data extracted from clinical trials included within the meta-analysis for hypertension

Study & design	Losartan				Candesartan			
	Mean (max) dose	Baseline	SBP [95% CI]	DBP [95% CI]	Mean (max) dose	Baseline	SBP [95% CI]	DBP [95% CI]
Lacourcière Y et al (1999)**(177)** – Forced titration – week 8	100mg (from 50mg at week 4) [n=115]	153.0/100.2 mmHg	Peak -10.3mmHg [-12.6 to -8.0] Trough -8.2mmHg [-10.7 to -5.7]	Peak -7.7mmHg [-9.3 to -6.2] Trough -5.8mmHg [-7.5 to -4.1]	16mg (from 8mg at week 4) [n=116]	155.1/101.8 mmHg	Peak -14.5mmHg [-16.8 to -12.3] Trough -12.4mmHg [-14.8 to -10.0]	Peak -9.4mmHg [-10.9 to -7.9] Trough -8.2mmHg [-9.9 to -6.5]
Gradman AH et al (1999)**(175)** – Not forced titration – week 8	100mg (from 50mg at week 4) [n=170; 56% not forced]	154.1/100.5 mmHg	Peak -14.4mmHg [-16.8 to -12.1] Trough -10.0mmHg [-12.2 to -7.8]	Peak -9.6mmHg [-11.1 to -8.2] Trough -8.9mmHg [-10.1 to -7.6]	32mg (from 16mg at week 4) [n=162; 52% not forced]	152.9/100.3 mmHg	Peak -16.5mmHg [-18.8 to -14.1] Trough -11.9mmHg [-14.1 to -9.6]	Peak -12.6mmHg [-14.0 to -11.1] Trough -11.0mmHg [-12.3 to -9.8]
Bakris G et al (2001)**(174)** – CLAIM study – Forced titration – week 8	100mg (from 50mg at week 2) [n=332]	152.0/99.9 mmHg	Peak -12.6mmHg [not reported] Trough -9.8mmHg [not reported]	Peak -10.1mmHg [not reported] Trough -8.7mmHg [not reported]	32mg (from 16mg at week 2) [n=322]	152.6/100.1 mmHg	Peak -15.2mmHg [not reported] Trough -13.3mmHg [not reported]	Peak -11.6mmHg [no CI] Trough -10.9mmHg [no CI]
Vidt DG et al (2001)**(176)** – CLAIM II study – Forced titration – week 8	100mg (from 50mg at week 2) [n=304]	152.2/100.2 mmHg	Peak -12.0mmHg [not reported] Trough -10.1mmHg [not reported]	Peak -9.5mmHg [not reported] Trough -9.1mmHg [not reported]	32mg (from 16mg at week 2) [n=307]	153.6/100.4 mmHg	Peak -15.5mmHg [not reported] Trough -13.4mmHg [not reported]	Peak -12.9mmHg [no CI] Trough -10.5mmHg [no CI]

Study & design	Losartan				Candesartan			
	Mean (max) dose	Baseline	SBP [95% CI]	DBP [95% CI]	Mean (max) dose	Baseline	SBP [95% CI]	DBP [95% CI]
Andersson OK and Neldam S (1998)**(172)** – No titration – week 8 (low dose)	50mg [n=83]	168/104 mmHg	Peak -19mmHg [not reported] Trough -11.0mmHg [not reported]	-12mmHg [not reported] Trough -7.0mmHg [not reported]	8mg [n=82]	169/102 mmHg	Peak -16mmHg [not reported] Trough -14.0mmHg [not reported]	Peak -10mmHg [no CI] Trough -9.0mmHg [no CI]
Andersson OK and Neldam S (1998)**(172)** – No titration – week 8 (high dose)	50mg [n=83]	168/104 mmHg	Peak -19mmHg [no CI] Trough -11.0mmHg [no CI]	Peak -12mmHg [no CI] Trough -7.0mmHg [no CI]	16mg [n=84]	168/103 mmHg	Peak -20mmHg [no CI] Trough -14.0mmHg [no CI]	Peak -12mmHg [no CI] Trough -10.0mmHg [no CI]
Baguet et al (2006)**(173)** – No titration – week 6	50mg [n=89]	161/101 mmHg	Trough -8.8mmHg [no CI]	Trough -5.1mmHg [no CI]	8mg [n=87]	160/101 mmHg	Trough -10.8mmHg [no CI]	Trough -7.3mmHg [no CI]
Willemsen et al (2004)**(179)** – No titration – week 4	50mg [n=4]	168/105 mmHg	Trough -23mmHg [no CI]	Trough -16mmHg [no CI]	16mg [n=4]	168/105 mmHg	Trough -30mmHg [no CI]	Trough -18mmHg [no CI]
Manolis AJ et al (2000)**(178)** – Not forced titration – week 12	100mg (from 50mg at week 6) [n=449; 47% had dose increased]	153.0/101.6 mmHg	Trough -14.4mmHg SD=11.7 [-15.5 to -13.3]	Trough -12.4mmHg SD=7.2 [-13.0 to -11.7]	16mg (from 8mg at week 6) [n=462; 45% had dose increased]	152.7/101.6 mmHg	Trough -15.8mmHg SD=12.2 [-16.9 to -14.7]	Trough -13.1mmHg SD=7.6 [-13.8 to -12.4]

5.6.4 Sensitivity analyses

To determine an accurate understanding of the potential ICERs, MS assisted in the modelling of four alternative scenarios using one-way sensitivity analyses from the base-case model (see Table 15).

Variation in baseline risk

The baseline risk was varied by increasing the cohort pre-treatment SBP in the range 140-180mmHg. The cost-effectiveness of candesartan decreased as the baseline risk lowered, as would be expected as the value of treatment reduces when fewer patients develop disease (see Figure 35). The range of ICERs is £41,469 to £52,644 for male patients and £41,591 to £85,244 for female patients.

Variation in hypertensive effectiveness

The cost-effectiveness variations from the base-case for male and female patients at the limits of the 95% confidence intervals for the trough SBP were identified using a random effects meta-analysis (-2.22 to -3.79 mmHg). The ICERs were £71,049 and £87,015 for a 2.22 mmHg SBP reduction in male and female patients, respectively, and £44,930 and £53,804 for a 3.79 mmHg SBP reduction in male and female patients, respectively. Therefore, within the identified pooled range of statistical uncertainty, losartan remained the most cost-effective treatment strategy. Figure 36 shows variation of the ICER with a wider range of trough SBP reduction (treatment effectiveness difference) with all other base case parameters left unchanged. For the ICER to fall below the UK perceived threshold of cost-effectiveness (£30,000 per QALY), the trough SBP differential between candesartan and losartan would have to be greater than 5 mmHg (in favour of candesartan).

Projected pricing for losartan

Using previous experience of generic market prices following patent expiry, it was anticipated that the acquisition cost of generic losartan may drop to £0.88 per pack of 28 tablets. Based on this figure, the base-case ICER would further increase to £74,901 and £91,368 for male and female patients, respectively, in favour of losartan. At this

generic price, candesartan becomes increasingly unfavourable from a cost-effectiveness point of view whilst it remains under patent protection and holds its current price (see Section 5.7.7 for further detail post candesartan patent expiry).

Variation in age

Variation in the cost-effectiveness with changes in the cohort starting age for male and female patients was also modelled. The ICER at 35 years of age was £151,140 and £369,075 for male and female patients, respectively, decreasing to the base-case values at age 65. Figure 37 shows variation in ICER with starting age of the cohort in the range 35-65 years for males and females. Throughout the wide range of initial ages, losartan dominated candesartan establishing its position as the most cost-effective intervention of the two.

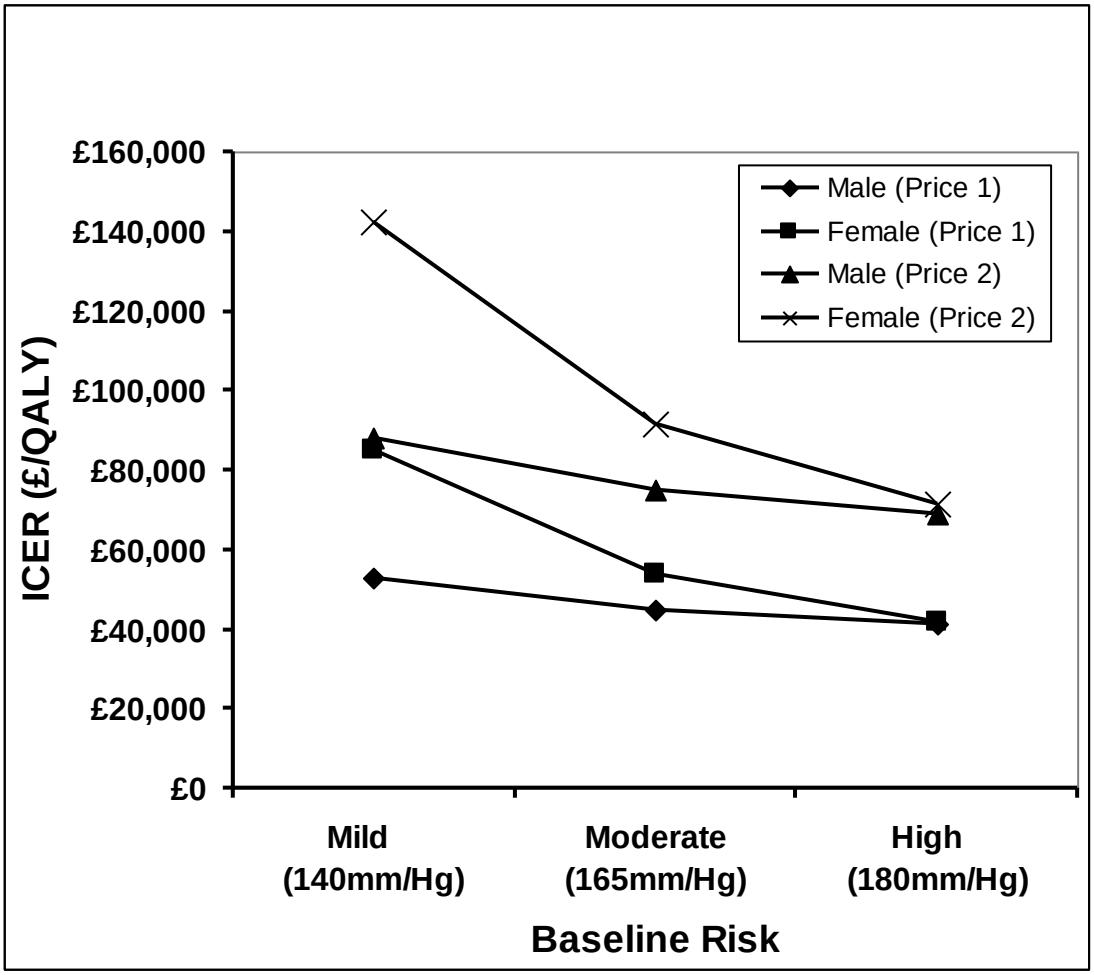


Figure 35: Incremental cost-effectiveness ratios (ICERs) of candesartan compared with losartan versus baseline hypertension risk.

Results are displayed for males and females at two generic losartan prices: Price 1 (£77.64 per annum) and Price 2 (£10.56 per annum). At all risk levels and for both genders the ICER is greater than £30,000 per QALY indicating that losartan is dominant (i.e. the more cost-effective option)

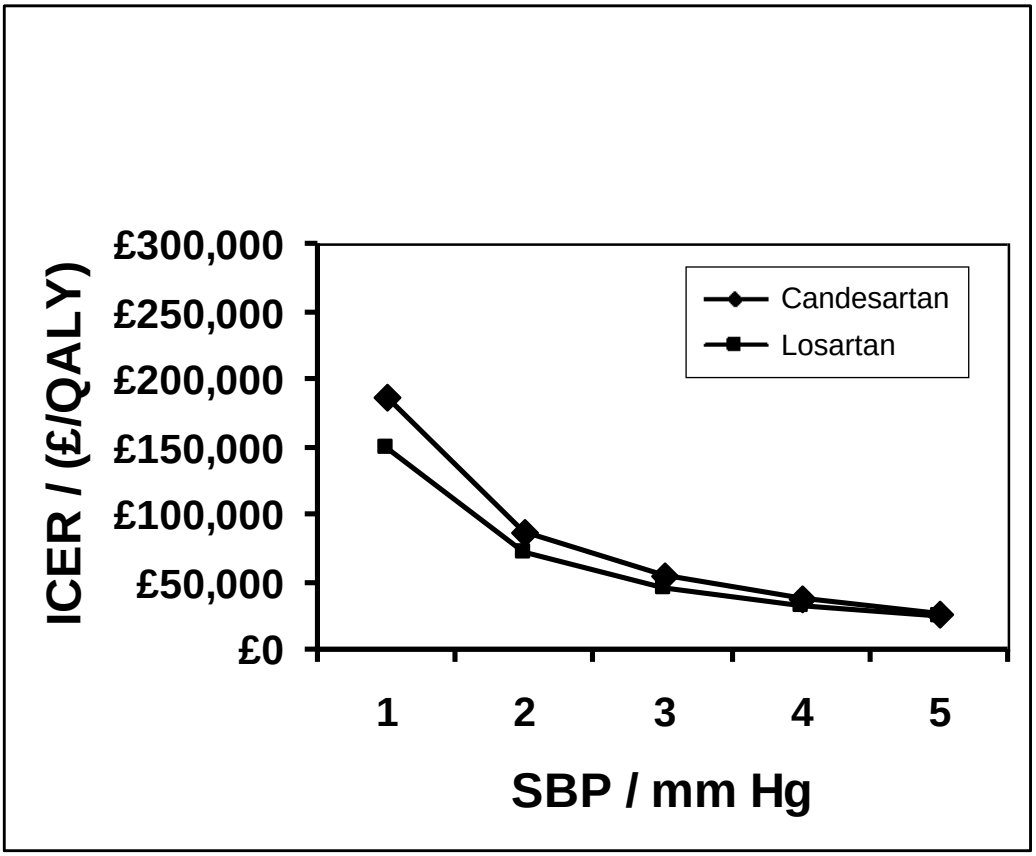


Figure 36: One-way sensitivity analysis on the incremental cost-effectiveness ratio (ICER) of candesartan relative to losartan versus difference in relative anti-hypertensive effect on systolic blood pressure (SBP) of candesartan relative to losartan.

Within the range 0-5 mmHg difference between the effectiveness (reduction in SBP) of both agents the ICER is greater than £30,000 per QALY indicating that losartan is dominant (i.e. the more cost-effective option)

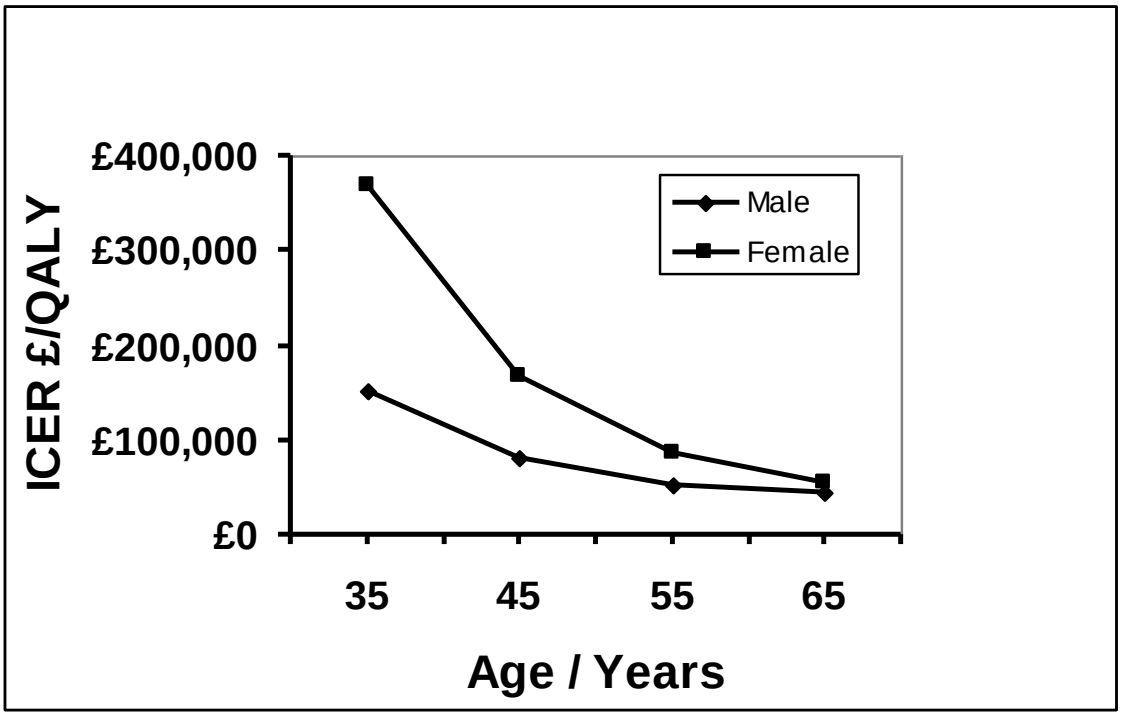


Figure 37: One-way sensitivity analysis on the incremental cost-effectiveness ratio (ICER) of candesartan relative to losartan versus cohort starting age in the age range 35-65 years for males and females.

The curves demonstrate losartan as being dominant across all age ranges (i.e. losartan is cost-effective irrespective of starting age) in accordance with the UK perceived threshold of cost-effectiveness at £30,000 per QALY

5.7 Summary

5.7.1 Hypertension

This analysis provides a comprehensive overall estimate of the between-treatment difference in efficacy (reduction in BP) for candesartan as compared with losartan. A between-treatment difference of -1.96 mmHg (95% CI -2.40 to -1.51) for trough DBP and -3.00 mmHg (95% CI -3.79 to -2.22) for trough SBP in favour of candesartan was observed in comparative studies at their maximal licensed doses.

Based on the calculated BP-differential and using a 10-year Markov model, it was estimated that the cost per QALY gained would exceed £40,000 if candesartan was used in preference to losartan. The base-case model used a cohort of patients with moderate CHD / CVD risk (baseline SBP of 165 mmHg) and the results demonstrate cost-effectiveness of losartan with ICERs [for candesartan] of £44,930 and £53,804, for male and female patients, respectively.

Sensitivity analyses suggest that the case for losartan adoption holds across plausible variation in CHD and CVD risk, relative hypertension reducing effect, gender and patient cohort starting age. This case is strengthened further in a situation where the acquisition cost of the generic drug falls further, as is the case once a number of generic manufacturers come to market creating a competitive market, or if the patient is female, younger, or mildly hypertensive. In addition, the population level benefit achieved by reducing BP could be attained by using cheaper low-dose combination therapy regimens.

5.7.2 Heart Failure

In the absence of comparative trials directly comparing the efficacy of candesartan versus losartan in CHF, a qualitative analysis of the key studies was undertaken as an alternative to a quantitative meta-analysis.

Candesartan

The CHARM-Alternative study (180) was a placebo-controlled randomised trial of candesartan in 2028 patients with a LVEF ≤ 40% who were intolerant of an ACE inhibitor (72% cough, 13% symptomatic hypotension, 12% renal dysfunction). The primary composite end-point was significantly reduced with candesartan (334／1,013) in comparison with placebo (406／1,015) [hazard ratio (HR) 0.77, 95% CI 0.67–0.89; p = 0.0004]. This corresponds to a relative risk reduction (RRR) of 23% (absolute difference = 7%) and a number needed to treat (NNT) of 14 (i.e. 14 patients need to be treated for the duration of the study to prevent one patient from dying of a cardiovascular event or being hospitalised for treatment of heart failure). The primary end-point was powered to a statistically significant level by the reduction in hospital admission for CHF [HR 0.68, 95% CI 0.57–0.81; p < 0.0001] as the reduction in cardiovascular death was non-significant [HR 0.85, 95% CI 0.71–1.02; p = 0.072].

The CHARM-Added study (189) (ACE-Inhibitor + candesartan versus ACE-Inhibitor + placebo) recruited 2,548 patients with an LVEF ≤ 40% who were receiving an optimal-tolerated dose of an ACE-inhibitor. The primary composite end-point (cardiovascular death and hospital admission for HF) was significantly reduced with candesartan (483／1,276) in comparison with placebo (538／1,272) (HR 0.85, 95% CI 0.75–0.96, p = 0.011). This corresponds to an RRR of 16% (absolute difference = 4.4%) and an NNT of 23. A statistically significant reduction was observed for both components of the primary endpoint: cardiovascular death [HR 0.84, 95% CI 0.72–0.98; p = 0.029] and hospital admission for CHF [HR 0.83, 95% CI 0.71–0.96; p = 0.014]. The composite secondary end-point of all-cause mortality or first CHF hospitalisation was also significantly reduced in the candesartan group [HR 0.87, 95% CI 0.78–0.98; p = 0.021]. Of note, at baseline, 55% of patients randomised to receive dual therapy were taking a beta-blocker, whilst only 17% were also taking spironolactone [aldosterone antagonist]. With regards to safety, of the 74 patients treated with candesartan + ACE inhibitor who were also taking spironolactone, three (4%) developed serum potassium levels > 6 mmol／l compared

with one of 71 (1%) in the placebo group (number needed to harm = 33).

Losartan

In comparison, the recently published HEAAL study (183) [low-dose losartan (50 mg daily) vs. high-dose losartan (150 mg daily)] recruited 3,846 patients with an LVEF ≤ 40% and a documented intolerance to an ACE inhibitor. The proportion of patients who met the primary composite end-point (death and admission for HF) was 43% (losartan 150 mg) versus 46% (losartan 50 mg), which was regarded as a modest yet significant benefit (HR 0.90, 95% confidence intervals 0.82–0.99; p = 0.027). In comparison with the CHARM-Added study, at baseline, 72% of patients randomised to both arms were taking a beta-blocker, and 38% were also taking an aldosterone antagonist. Overall, the authors reported superiority of losartan 150 mg once daily over 50 mg once daily for the treatment of CHF. With regards to safety, losartan 150 mg daily compared with 50 mg daily was noted to cause a significant reduction in glomerular filtration rate (6.1 ml／min vs. 1.9 ml／min; p < 0.001, respectively). The incidence of premature discontinuation from therapy as a result of hyperkalaemia, hypotension, renal impairment and angioedema was non-significantly different between the two arms (p = 0.20, p = 0.65, p = 0.22, and p = 0.12, respectively). The investigators put forward the hypothesis that up-titrating the dose of losartan as monotherapy may provide equally favourable results to a combination of ACE inhibitor and ARB as demonstrated in the CHARM-Added study.

UK Licensing

Both losartan and candesartan hold a marketing authorisation for the treatment of essential hypertension and CHF (in patient with an LVEF ≤ 40%). Two key differences in their licensing criteria are noted:

i. Essential hypertension: candesartan is indicated for adults only, whereas losartan is also indicated for children and adolescents aged 6–18 years; and
ii. CHF: candesartan is indicated for use as monotherapy or in combination with an ACE inhibitor (following disease progression, incompatibility, or contraindication), whereas

losartan is only indicated as monotherapy in place of an ACE inhibitor.

5.7.3 Strengths and Limitations

The use of a trade-off analysis in estimating the comparative cost-effectiveness of these two AIIRAs has several strengths. First, all data included within this analysis were extracted from robust trials which met strict inclusion criteria. The requirement of a randomised, double-blinded, controlled trial ensured that the final dataset would be subject to the least possible amount of bias. Second, the efficacy outcome selected (reduction in SBP) is well accepted as an informative outcome measure concerning hypertension and risk factor for the development of CVD. Third, all included study manuscripts were published in full allowing for intention to treat analyses.

However, the results obtained from this methodology are subject to the limitations that are inherent in any meta-analysis. First, individual prospective studies only provide information over a short period of time (4–12 weeks) whilst the implications are extrapolated for life-long therapy. Second, excluding trials that have not been published may exaggerate the treatment effects observed as publications tend to favour those with positive results. Third, pooling of data from trials with differences in trial design, methodology and patient groups may result in a heterogeneous dataset from which conclusions are drawn. However, such differences in patient groups may serve to strengthen the meta-analysis by allowing generalisability of the results to a broader group. Furthermore, the use of a random-effects model and tests to identify the presence of significant heterogeneity aid to minimise and highlight the impact of such effect. Fourth, this analysis was restricted to data relating to the adult population; therefore it cannot be directly extrapolated to patients under the age of 18 years, although this is not a common target population for the management of hypertension or heart failure. Fifth, doses employed within the clinical trials are not always consistent with those used in clinical practice, therefore limiting the external validity of these data. To minimise the extent of this, only data comparing high-dose losartan

with high-dose candesartan were used in the subsequent CUA. Finally, data within this analysis only relates to the use of losartan and candesartan for the treatment of hypertension in patients with no other comorbidities and therefore should be extrapolated with caution in such patients.

To ensure a robust analysis, best evidence was used, including the well validated Framingham risk equations as criteria for the development of disease and subsequent mortality in the CUA. The model was developed by an experienced health economists, however, is subject to some caveats. The model calculates only the first episode of MI or stroke events and the subsequent quality of life and costs associated with survivors. It does not consider that stroke patients may experience a fatal MI or vice versa. Whilst a fully probabilistic model may have been constructed, for ease of understanding and presentation it was decided to use deterministic analyses allowing for uncertainty using sensitivity analysis. This showed, that under plausible variation in key parameter values using one-way sensitivity analyses, that losartan remains the more cost-effective of the two studied AIIRAs.

5.7.4 Place in therapy for AIIRAs

Currently, there are 11 ACE inhibitors and seven AIIRAs available in the UK (186). The UK NICE offers guidance on the use of these drugs in the following areas: CHF, hypertension, MI (secondary prevention), type II diabetes and chronic kidney disease.

Where either an ACE inhibitor or an AIIRA is indicated, NICE recommends that an ACE inhibitor is routinely the drug of choice on the basis that there is a more robust evidence-base for their use. A recent review also suggested that AIIRAs may be less effective than even cheaper ACE inhibitors for MI protection (190). However, a meta-analysis has suggested that AIIRAs are indeed as effective as ACE inhibitors on the risk of MI, cardiovascular mortality and total mortality and also concluded that they may even be slightly more protective than ACE inhibitors on the risk of stroke (191). As CHD is more common than cerebrovascular disease it seems reasonable that

ACE inhibitors remain first-line in all (except perhaps in some high-risk groups).

For the treatment of hypertension, if either an ACE inhibitor or an AIIRA is indicated, the NICE recommendation stipulates using a drug that can be taken once-daily, is generically prescribed and minimises cost (188). For heart failure, combined treatment with both an ACE inhibitor and an AIIRA currently has a limited role. For patients with heart failure who remain symptomatic despite the use of a diuretic, beta-blocker and ACE inhibitor, options include to either add in an aldosterone antagonist, an AIIRA or hydralazine combined with a nitrate (192). Routine treatment with an ACE inhibitor in combination with an AIIRA is not recommended as recent publications have found little evidence to support this (193, 194). A recently published meta-analysis regarding the use of an AIIRA in combination with an ACE inhibitor concluded that there was no clear survival benefits associated with the combination treatment strategy.(195)

5.7.5 AIIRA prescribing recommendations

The recommendation from this analysis is that generic losartan be initiated in all patients indicated for an AIIRA in both hypertension and heart failure. AIIRAs should not routinely be combined with ACE inhibitors and primarily be used in patients who are ACE-inhibitor intolerant. For existing patients on candesartan for hypertension it is recommended that patients are changed to losartan except in the rare scenario of prior intolerance to losartan. For existing patients on candesartan for heart failure where not already on maximal target dose, AIIRA dose escalation is encouraged. If a patient on candesartan is due for dose escalation, it is recommended that changing to losartan is considered at this point unless intolerant. Where a patient is on maximal dose candesartan for heart failure, it is recommended that the decision to switch be considered on a case-by-case basis by the responsible physician.

5.7.6 Dosing

The losartan dose used in hypertension should be 50mg or 100mg daily depending on whether the existing AIIRA dose is at the lower or

upper end of its dosing schedule. A dose of 150mg daily is recommended as the target dose for heart failure; although this is currently outside of the manufacturers Marketing Authorisation (off-label), it is supported by clinical data.

5.7.7 Post candesartan patent expiry

The patent protection for candesartan is in place until December 2012. Following this point, it is highly likely that a number of generic manufacturers will launch their version of generic candesartan against a similar pricing strategy to that seen with generic losartan. Given that it is common practice for patent holders to submit requests for patent extensions and typically receiving a 6 month extension, the anticipated UK availability date for generic candesartan would be circa summer 2013. As generic losartan has been on the market since 2011 it will initially still represent the lowest price and most cost-effective treatment option for hypertension and heart failure, however over a period of 12 month the prices of the two treatments will likely become comparable. In this situation, the cost utility analysis described above would become obsolete with prescribing favouring candesartan owing to its greater reduction in blood pressure.

5.8 Conclusion

Although candesartan, the most widely prescribed AIIRA, reduces both SBP and DBP to a greater extent when compared with losartan, this does not translate to a cost-effective reduction based on an assessment of its efficacy (using the validated Framingham model) compared with current and near future acquisition costs of losartan alongside perceived NHS affordability thresholds. No robust evidence supporting the superiority of candesartan over losartan in the treatment of heart failure was found. A recommendation of the analysis is therefore to adopt generic losartan as the AIIRA of choice, for new and existing patients, which could, based on 2009 prescribing figures in primary care alone, save the UK NHS approximately £200 million per annum. This saving may be enhanced as the acquisition price of generic losartan reduces over time, and further expanded following the availability of generic candesartan at a similar price to generic losartan.

This work has been published in the International Journal of Clinical Practice 2011; 65(3): 253-263 (doi: 10.1111/j.1742-1241.2011.02633.x) and meets the criteria for inclusion on the NIHR Centre for Reviews and Dissemination (CRD) Database of Abstracts of Reviews and Effects (DARE).

This work was also identified by the NHS Economic Evaluation Database and cited on their website, reference CRD/NHSEED/22011000527, citing the following comments:

- A justification for the selection of comparators was given
- The selection of clinical data was well carried out
- The cost categories were consistent with the perspective and data sources reflecting the UK NHS setting
- An appropriate incremental approach was used to synthesis the costs and benefits
- The cost-effectiveness framework was conventional and the clinical analysis was very well conducted
- The authors conclusions appear to be robust

5.9 Implications of Research

The findings from this area of research were published, communicated locally and circulated nationally. Local guidelines, clinician and patient letters were written to support prescribing and treatment switching. See section 10.3.2 to 10.3.5 for further details. See section 5.9.1 for implications on prescribing and spend.

5.9.1 Implications on prescribing and spend

Further to the publication of the above finding in the International Journal of Clinical Practice (March 2011), the circulation of above letters to clinicians and patients permitted the actionable roll-out of the results of the research project. Using data publically available from the NHS Business Services Authority (Prescription Services Division) as hosted on www.data.gov.uk it is possible to track the monthly spend on each medicine line by GP practice within the UK.

The change in prescribing illustrations below highlights: (1) the effect of the publication and resultant shift in AIIRA agent being prescribed to losartan, from 8,000 items per quarter in March 2011 to 14,000 items per quarter in June 2012 (see Figure 38); and (2) the consequence of this shift in prescribing on spend upon availability of a generic medicine within the market place, with spend on other ARBs reducing from £120,000 per quarter in March 2011 to £40,000 per quarter in June 2012 (see Figure 39). Following the loss of patent protection for candesartan it would be expected that a number of generic manufacturers will launch their version of generic candesartan against a similar pricing strategy to that seen with generic losartan. This will immediately result in a reduction in spend (NIC) as per losartan in Figure 39. Given the favourable efficacy profile observed with candesartan compared with losartan, in the scenario where both are similarly priced as generics it is anticipated that prescribing will revert back to generic candesartan in the higher risk patients and generic candesartan placed as first-line for new patients where an AIIRA is indicated.

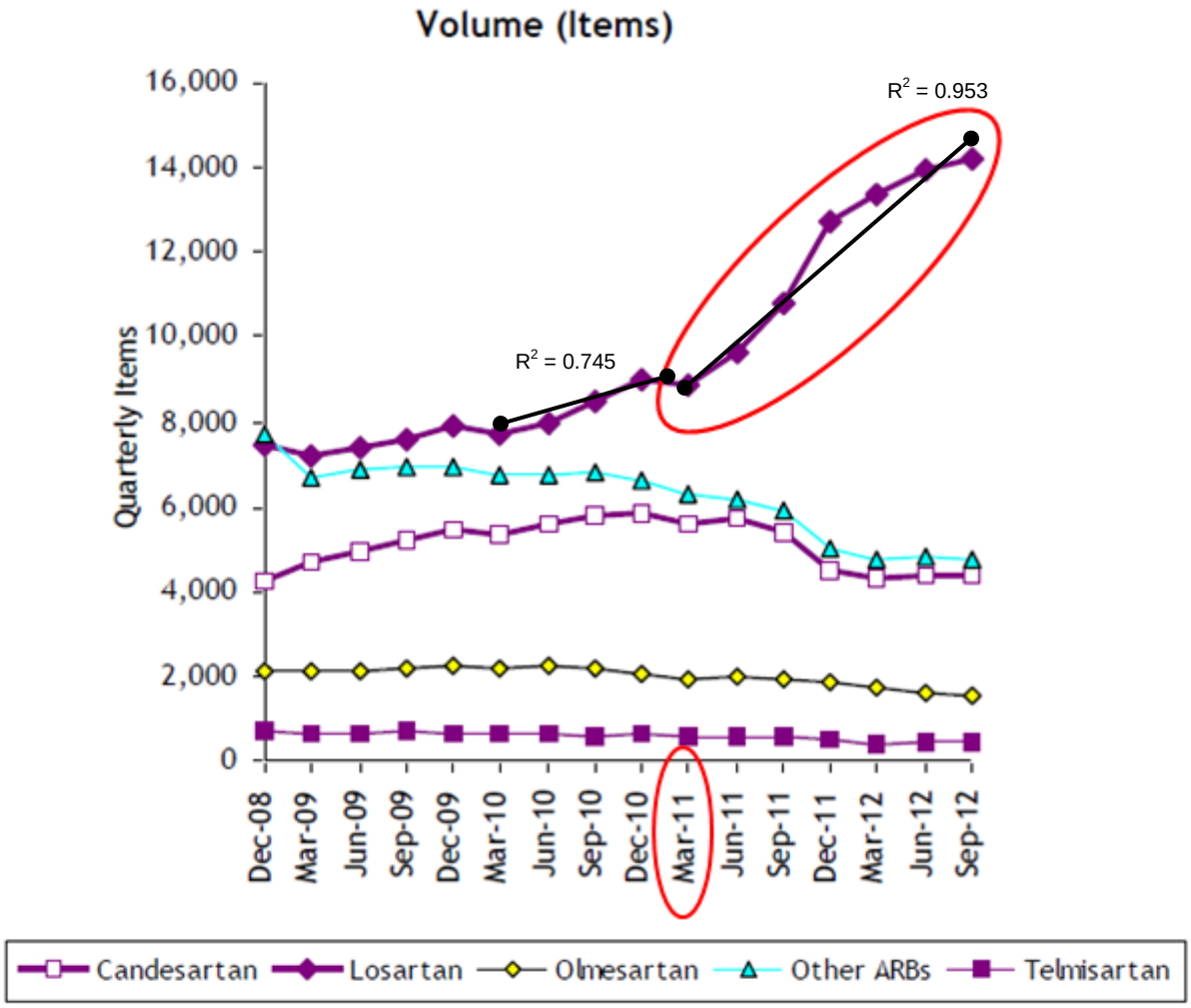


Figure 38: Volume (number of items) of angiotensin-receptor blockers (ARBs) issued by General Practitioners and other primary care prescribers in England and Wales (2008-2012).

The highlighted date of March 2011 indicates publication of the analysis and findings, which correlates with a change in prescribing trend (i.e. the number of items of losartan doubles over a period of 12 months whilst other ARBs decline). The R^2 values are indicative of the linear trendline March 2010-March 2011 vs. March 2011-September 2012.

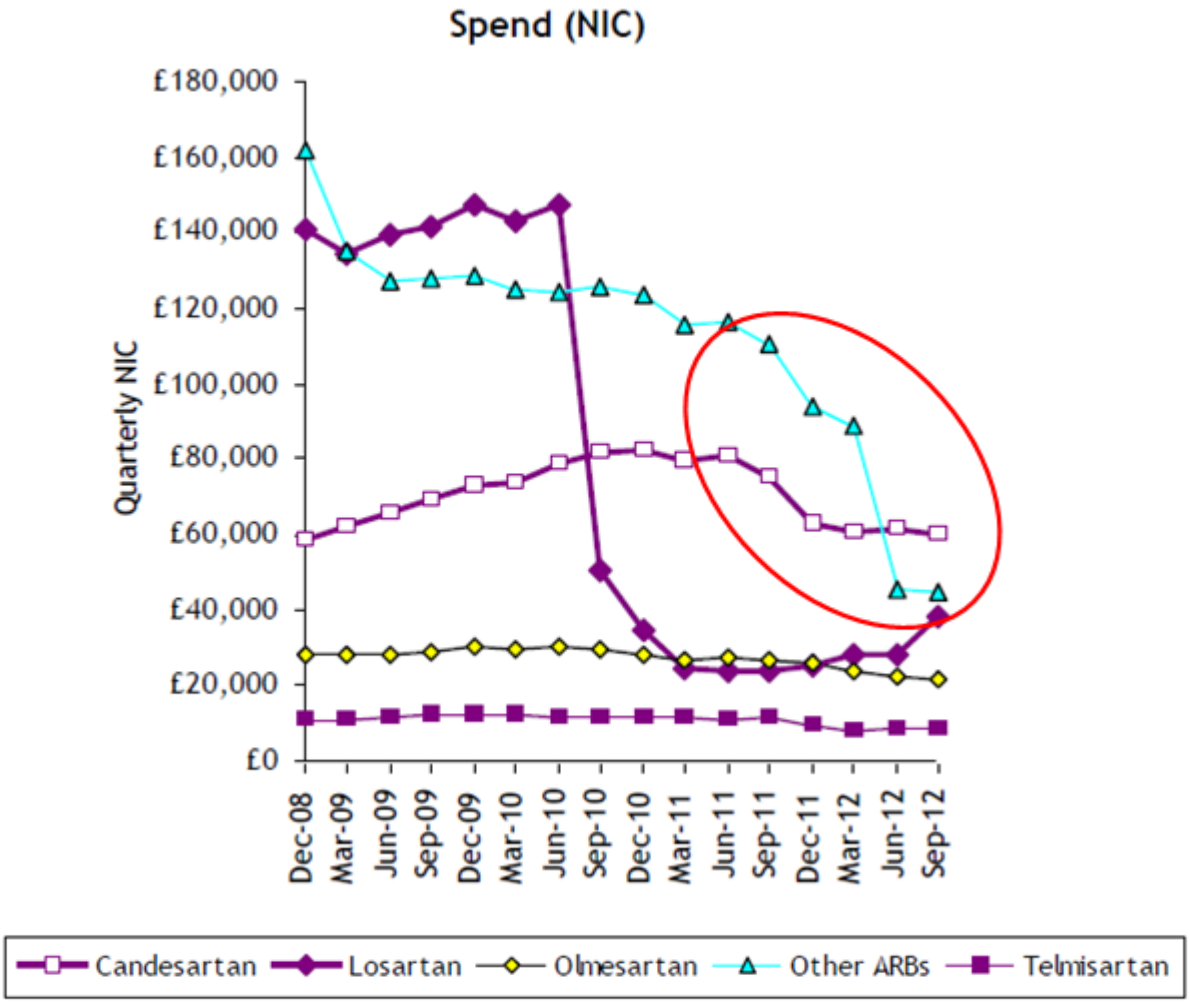


Figure 39: Spend (net ingredient cost) of angiotensin-receptor blockers (ARBs) issued by General Practitioners and other primary care prescribers 2008-2012.

The highlighted area indicates the rapid decline in spend on candesartan and other ARBs post publication of the analysis and findings.

[BLANK PAGE]

6.0 Chapter 6: Research Project Three – Antimuscarinics and newer agents for the Management of Overactive Bladder syndrome

6.1 Introduction

This third research project reintroduces the use of Bayesian NMA and continues with the trade-off methodology based on multiple efficacy and tolerability endpoints. Between completion of the first research project and commencement of this third project, the Bayesian statistical technique has become adopted by NICE although traditionally used for simpler comparisons. The Bayesian NMA will be undertaken to comparatively assess data from all published RCTs comparing the antimuscarinics for overactive bladder meeting strict criteria, including head-to-head design where available, in order to obtain estimates on their relative efficacy and safety. Certain studies were conducted with a design to meet regulatory criteria whilst also attempting to be clinically useful, i.e. comparison to placebo and introducing an active comparator. The use of a Bayesian analysis rather than the frequentist analysis is therefore suitable as a number of published three-arm trials have investigated these agents in a direct head-to-head comparison thus permitting the construction of a network treatment loop. This design will require the coding for the Bayesian analysis to be completely revised to that used in the AED research project.

The goal for this project is a hybrid of the first and second projects, to utilise the Bayesian NMA supported by the trade-off analysis in order to create a hierarchy for the range of medicines within this specific class. This hierarchy would then be used to influence prescribing at the local specialist centres as well as across the local health economy. To ensure a robust and clinically relevant analysis based on learning from project one, consultation with a methodological and clinical expert was undertaken.

6.2 Literature Review

6.2.1 Overactive Bladder syndrome

Overactive bladder (OAB) is defined by the International Continence Society as urinary urgency, usually accompanied by frequency and nocturia, with or without urgency urinary incontinence (UUI), in the absence of urinary tract infection (UTI) or other obvious pathology.(196) This is a widespread, chronic illness that affects the lives of millions of people worldwide at all ages. The incidence increases with age, with a prevalence of up to 31% in women and 42% in men > 75 years.(197) OAB has a major impact on quality of life,(198) as well as imposing a substantial socio-economic burden, with direct annual costs comparable to that of other chronic diseases such as diabetes mellitus.(199)

Several treatment options are available for OAB including bladder and behavioural training, pharmacologic treatment and surgical therapies.(200) Oral antimuscarinics represent the mainstay of pharmacologic treatment for the management of OAB. They are recognised to be effective in the improvement of OAB symptoms and have a good safety profile.(201)

6.3 Quantitative Analysis

6.3.1 Current practice guidance

The utilisation of systematic reviews and meta-analysis of the urological literature is proposed as a crucial step in improving the quality of urological patient care(202) with decisions based on a balance of benefits, potential harm and costs.

Despite the publication of a NICE Clinical Guideline and several systematic reviews and meta-analyses of pharmacological treatment for OAB within the last few years, no clear differences in efficacy between antimuscarinics were found.(203-205)

The aim of this quantitative trade-off analysis is to compare the clinical efficacy and tolerability of the most widely used pharmacologic treatments for OAB, and to determine an overall hierarchy to guide clinicians.

6.4 Collaboration

All work undertaken within this chapter of the thesis was undertaken by me. Systematic review support was provided by Dr Grosso in the form of a second independent check against the proforma. Advice on the appropriate use of WinBUGS coding and analysis of findings was provided by Dr Welton with all analyses performed by me.

6.5 Protocol

6.5.1 Search strategy

The literature search was undertaken according to the guidelines of the Centre for Reviews and Dissemination and Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA)(206) statement further to the publication by Maman K et al.(205) To identify randomised controlled trials (RCTs) analysing the efficacy and safety of pharmacologic treatments for OAB, a systematic literature search was conducted using Medline (Ovid), Embase (Ovid) and the Cochrane Central Register of Controlled Trials on 08 June 2014. Refer to Table 17, Table 18 and Table 19, respectively for a full list of search terms.

6.5.2 Eligibility criteria

This review considered all published RCTs which investigated the efficacy and safety of pharmacological treatments in the management of OAB. Case reports, case series, and database studies were excluded from the review. Although the use of language restriction was not found to bias outcomes (as discussed in section 4.6.5), no language restriction or geographical restriction were applied here. All publications from 2000 onwards were included in the original search by Maman et al which was supplemented to 08 June 2014 for inclusion in the search. Only fully published articles were included. Letters, abstracts, conference posters, literature reviews and unpublished manuscripts were excluded.

Reviewed studies enrolled adults (≥ 18 years of age) diagnosed with OAB. Studies among patients with neurogenic detrusor overactivity and men with lower urinary tract symptoms associated with benign prostatic hyperplasia were excluded. Further exclusions were studies

investigating pharmacological treatment in a single sex population or which incorporated a surgical intervention. Refer to Table 20 for full list of inclusion and exclusion criteria for studies eligible for the mixed-treatment comparison (MTC) analysis.

This review included studies on medicines licensed for the management of OAB within the UK at doses within their Marketing Authorisation which fulfilled the above inclusion and exclusion criteria.

6.5.3 Outcome measures

To be eligible, a study had to report a measure of efficacy and / or safety of OAB treatment over a minimum of 12 weeks. The efficacy endpoints, specified *a priori*, were (1) micturition frequency and (2) urgency urinary incontinence (UUI) episodes, both as per 24 hours. Studies that reported their results as weekly episodes were converted to an average number of episodes per 24 hours. The safety endpoints, specified *a priori*, were the number of patients reporting (1) dry mouth and (2) constipation, over the treatment period. These endpoints were validated by a Consultant within this field to ensure clinical relevance.

6.5.4 Study selection

Publications identified through the electronic searches were assessed independently for relevance by me and validated by Dr Grosso. Any disagreement between reviewers was resolved by discussion and consensus. In a second step, the reviewers read the full text of the retrieved references and selected the articles that met the inclusion criteria. The articles finally selected for the review were checked to identify different articles related to the same study. The selection process was recorded in accordance with the PRISMA flowchart process (see Figure 40).

6.5.5 Quality assessment and assessment of bias

Abstracts and full text articles were retrieved from the search, with potentially relevant publications selected against the pre-specified inclusion and exclusion criteria. To ensure consistency of data

extraction for each study, a structured form was used. The articles were retrieved by me and verified by Dr Grosso. The risk of bias assessment was also undertaken by me and then independently verified by Dr Grosso using criteria associated with sequence generation, allocation concealment, baseline comparability, blinding, follow-up, and selective reporting, as recommended by the Cochrane Collaboration (Chapter 8: Assessing risk of bias in included studies).(36) Disagreements between reviewers were resolved by discussion and consensus.

Table 17: Search strategy used in MEDLINE (Ovid) for the analysis of antimuscarinics for the management of OAB syndrome (search date range: 2013 week 24 to current; searched on 08 June 2014)

ID	Searches	Results
1	exp Urinary Bladder, Overactive/	2403
2	exp Urinary Incontinence/	26156
3	exp Urinary Incontinence, Urge/	540
4	or/1-3	27911
5	exp Muscarinic Antagonists/	49017
6	(solifenacin or Vesicare or Vesikur or Vesiker or Vesitirim).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	360
7	(tolterodine or Detrusitol or Detrusitol XL or Detrol or Detrol LA).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	813
8	(mirabegron or YM-178 or Betanis or betmiga).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	103
9	(darifenacin or Enablex or Emselex).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	286
10	(fesoterodine or Toviaz).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	137
11	(oxybutynin or Ditropan or Lyrinel XL).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	1257
12	(propiverine or Detrunorm).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	268
13	(trospium or Regurin or Flotros or Sanctura or Tropez or Trosec or Spasmex).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	234
14	or/5-13	50411
15	Randomized controlled trials as Topic/ or randomized controlled trial.mp	103516
16	Random allocation/ or random allocation.mp	81879
17	Or/15-16	184540
18	4 and 14 and 17	131
19	Limit 18 to (humans)	130
20	Limit 19 (2013 week 1 to current)	8

Table 18: Search strategy used in EMBASE (Ovid) for the analysis of antimuscarinics for the management of OAB syndrome (search date range: 2013 week 24 to current; searched on 08 June 2014)

ID	Searches	Results
1	exp overactive bladder/	8561
2	exp urge incontinence/	4642
3	exp urinary urgency/	3767
4	Exp urine incontinence/	51599
5	Or/1-4	57827
6	exp muscarinic receptor blocking agent/	56069
7	(solifenacin or Vesicare or Vesikur or Vesiker or Vesitirim).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	1352
8	(tolterodine or Detrusitol or Detrusitol XL or Detrol or Detrol LA).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	2993
9	(mirabegron or YM-178 or Betanis or betmiga).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	310
10	(darifenacin or Enablex or Emselex).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	1120
11	(fesoterodine or Toviaz).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	522
12	(oxybutynin or Ditropan or Lyrinel XL).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	4739
13	(propiverine or Detrunorm).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	1234
14	(trospium or Regurin or Flotros or Sanctura or Tropez or Trosec or Spasmex).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	1217
15	or/6-14	58983
16	randomized controlled trial/	343012
17	Random allocation.mp or randomization/	63302
18	Or/16-17	397197
19	5 and 15 and 18	501
20	Limit 19 to (humans)	489
20	Limit 20 (2013 week 1 to current)	54

Table 19: Search strategy used in The Cochrane Library for the analysis of antimuscarinics for the management of OAB syndrome (search date range: 2013 week 24 to current; searched on 08 June 2014)

ID	Search	Hits
#1	MeSH descriptor **Urinary Bladder, Overactive** explode all trees	314
#2	MeSH descriptor **Urinary Incontinence, Urge** explode all trees	81
#3	overactive near/3 bladder	992
#4	Urge near/3 incontinence	579
#5	(#1 OR #2 OR #3 OR #4)	1375
#6	MeSH descriptor **Muscarinic Antagonists** explode all trees	597
#7	solifenacin or Vesicare or Vesikur or Vesiker or Vesitirim	177
#8	tolterodine or Detrusitol or Detrusol XL or Detrol or Detrol LA	447
#9	mirabegron or YM-178 or Betanis	46
#10	darifenacin or Enablex or Emselex	71
#11	fesoterodine or Toviaz	114
#12	oxybutynin or Ditropan or Lyrinel XL	403
#13	propiverine or Detrunorm	96
#14	trospium or Regurin or Flotros or Sanctura or Tropez or Trosec or Spasmex	130
#15	(#6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14)	1439
#16	(#5 AND #15)	731
#17	Limit #16 (2013 week 1 to current; Trials; Cochrane Group)	103

Table 20: List of inclusion and exclusion criteria for the mixed-treatment analysis of antimuscarinics for the management of OAB syndrome

		Efficacy	Safety
Inclusion criteria	Publication status	Published	Published
	Study Design	Randomised Controlled trial, Double-blind	Randomised Controlled trial, Double-blind
	Patient population	OAB patients only	OAB patients only
	Results population	FAS, ITT	FAS, ITT, SAF
	Intervention	Studies comparing two or more of the following treatments: darifenacin, solifenacin, tolterodine, mirabegron, fesoterodine, oxybutynin, trospium or placebo.	Studies comparing two or more of the following treatments: darifenacin, solifenacin, tolterodine, mirabegron, fesoterodine, oxybutynin, trospium or placebo.
	Duration	Minimum 12 weeks	Minimum 12 weeks
	Outcomes	Independent endpoints 1. Change from baseline of the number of micturition / 24h 2. Change from baseline of the number of UUI / 24h	1. Number of patients experiencing dry mouth 2. Number of patients experiencing constipation
Exclusion criteria		1. Studies that are unpublished or in abstract / conference form only 2. Studies with less than two arm treatments 3. Studies comparing combined treatments 4. Flexible-dose studies 5. Studies with included a surgical paradigm 6. Studies vs. placebo patch 7. No standard error, standard deviation, variance or confidence interval around estimates of mean changes	1. Studies that are unpublished or in abstract / conference form only 2. Studies with less than two arm treatments 3. Studies comparing combined treatments 4. Flexible-dose studies 5. Studies with included a surgical paradigm 6. Studies vs. placebo patch

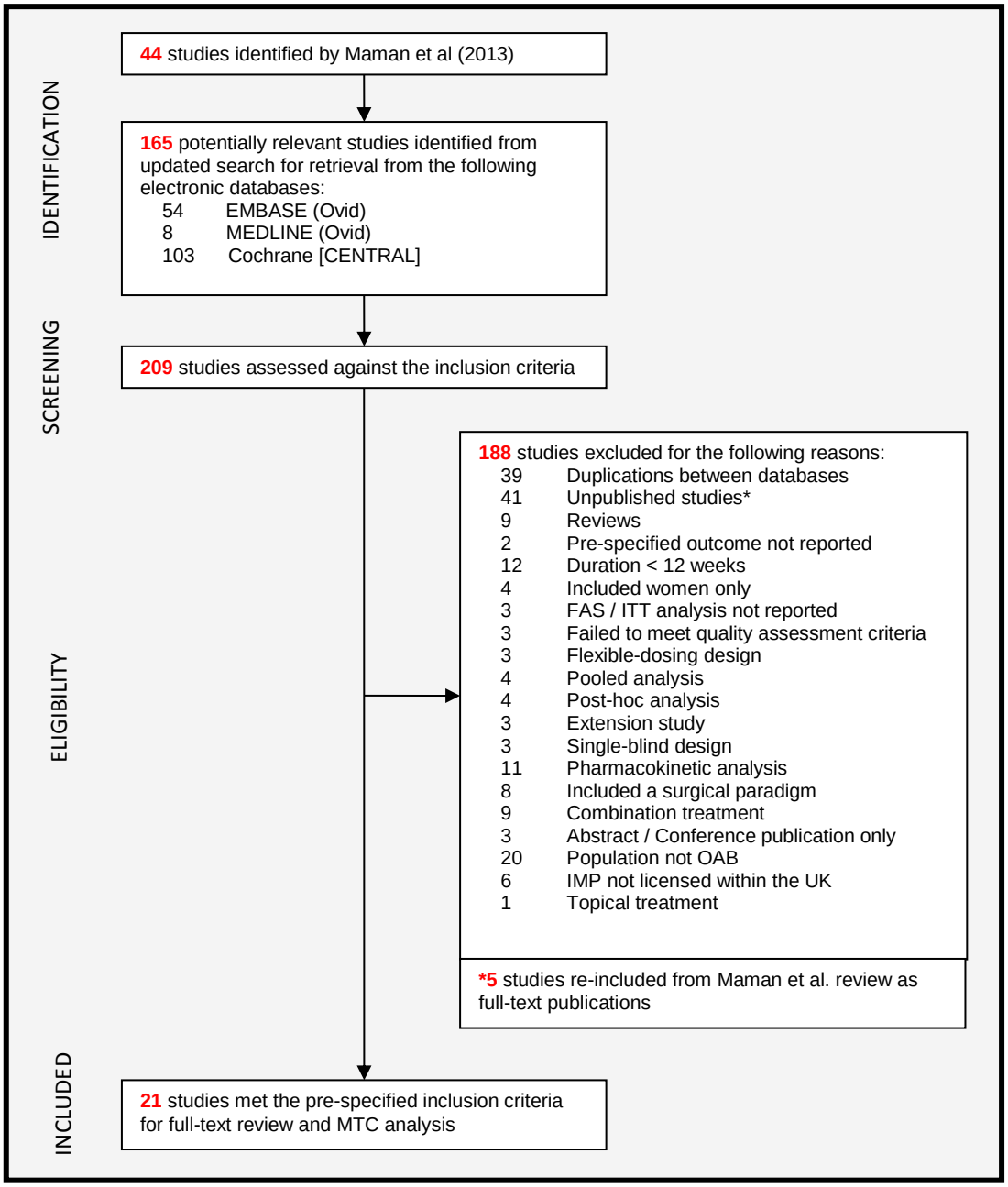


Figure 40: The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of studies identified from the systematic review for inclusion in the network meta-analysis for management of overactive bladder syndrome

6.5.6 Data analysis: Mixed Treatment Comparison (MTC)

A Bayesian mixed-treatment comparison (MTC) was conducted to estimate the relative efficacy and safety of all eligible pharmacological treatments compared with placebo, including all indirect comparisons. This was supplemented by a trade-off analysis of the pre-specified endpoints to establish an overall rank for each of the four parameters under investigation. Analyses were conducted for the general OAB population using reported full analysis dataset (FAS) or intention to treatment (ITT) population data. For each population, a random-effects model was used at the level of trials.

6.5.6.1 MTC code

The treatment effects from the MTC model were estimated using WinBUGS,(207) with the code adapted from the NICE Evidence Synthesis Technical Series Documents 2 and 3. (61-63) As the data were not normally distributed within the studies included, the MTC model was coded for binomial likelihood within a random effects model. The model for the efficacy outcomes is reflective of continuous data, whilst the model for the tolerability outcomes is reflective of dichotomous data. The median of the posterior distribution were taken as the point estimate and the 2.5th and 97.5th centiles provided the 95% credible interval (CrI). Refer to Figure 41 and Figure 42 for the WinBUGS codes created and used for the efficacy and tolerability outcomes, respectively.

6.5.6.2 MTC model

Unlike the code used for the AED chapter, a non-informative prior was selected for this MTC analysis. This was on the basis that no beliefs were in place prior to the analysis regarding the value of each effect size. Forty thousand iterations were used for each chain in the Bayesian analysis following a burn-in of 20,000. Convergence was assessed via visual assessment of trace plots. The visual assessment was corroborated through assessment of the ratio of the Monte Carlo error (MCe). Goodness of fit of the model was also assessed through analysis of the posterior mean of the sum of the residual deviance contributions of each data point.(62, 68)

6.5.6.3 Endpoints

For the efficacy endpoints, the treatment effect was the mean difference in the change from baseline for each treatment compared with placebo, under a binomial distribution for each arm. For the safety endpoint, the number of events reported in each treatment of a given trial was assumed to follow a binomial distribution with parameters n, the number of patients in a treatment arm, and p, the ‘‘true’’ probability of adverse events for each treatment. Placebo was used as the reference comparator for estimating the MTC models.

Posterior median values and 95% CrI are reported for mean differences in changes in symptoms from baseline to week 12 between treatments and odds ratios (ORs) for adverse events. A result was considered statistically significant when it had a probability of at least 97.5%. For example, a treatment was considered significantly more efficacious than placebo for reduction in number of micturition episodes per 24 hours if the probability of the difference in change in micturition frequency being negative was at least 97.5% (i.e. if the upper limit of the 95% CrI around the difference was less than zero).

Figure 41: WinBUGS model for the OAB network meta-analysis (efficacy endpoint)

```
# Binomial likelihood, identity link
# Randomeffects model for multi-arm trials

model{                                  #PROGRAM STARTS

for(i in 1:ns){                         #LOOP THROUGH STUDIES
    w[i,1] <- 0                         #Adjustment for multi-arm trials is zero for
control arm
    delta[i,1] <- 0                             #Treatment effect is zero for control
arm
     mu[i] ~ dnorm(0,.0001)             #Vague priors for all trial baselines

for (k in 1:na[i]) {                            #LOOP THROUGH ARMS
    var[i,k] <- pow(se[i,k],2)          #Calculate variances
    prec[i,k] <- 1/var[i,k]             #Set precisions
    y[i,k] ~ dnorm(theta[i,k],prec[i,k])    #Binomial likelihood
    theta[i,k] <- mu[i] + delta[i,k]    #Model for linear predictor

#Deviance contribution
      dev[i,k] <- (y[i,k]-theta[i,k])*(y[i,k]-theta[i,k])*prec[i,k]
     }

#Summed residual deviance contribution for this trial
   resdev[i] <- sum(dev[i,1:na[i]])
   for (k in 2:na[i]) {                 #LOOP THROUGH ARMS

#Trial-specific LOR distributions
      delta[i,k] ~ dnorm(md[i,k],taud[i,k])

#Mean of LOR distributions, with multi-arm trial correction
      md[i,k] <-  d[t[i,k]] - d[t[i,1]] + sw[i,k]

#Precision of LOR distributions (with multi-arm trial correction)
      taud[i,k] <- tau *2*(k-1)/k

#Adjustment, multi-arm RCTs
      w[i,k] <- (delta[i,k] - d[t[i,k]] + d[t[i,1]])

#Cumulative adjustment for multi-arm trials
      sw[i,k] <- sum(w[i,1:k-1])/(k-1)
    }
 }
totresdev <- sum(resdev[])              #Total Residual Deviance
d[1]<-0                                 #Treatment effect is zero for control arm

#Vague priors for treatment effects
for (k in 2:nt){  d[k] ~ dnorm(0,.0001) }
sd ~ dunif(0,5)                         #Vague prior for between-trial SD
tau <- pow(sd,-2)                       #Between-trial precision = (1/between-trial
variance)


# All pairwise comparisons Ranking
for (c in 1:(nt-1)) { for (k in (c+1):nt) { diff[c,k] <- (d[c] - d[k] )}}
for (k in 1:nt) {
rk[k] <- nt+1-rank(d[],k)               # assumes events are "good"
best[k] <- equals(rk[k],1)              #Calculate probability that treat k is best
}                                       #PROGRAM ENDS
```

Figure 42: WinBUGS model for the OAB network meta-analysis (tolerability endpoint)

```
# Binomial likelihood, logit link
# Randomeffects model for multi-arm trials

model{                                  # PROGRAM STARTS

for(i in 1:ns){                         # LOOP THROUGH STUDIES
  w[i,1] <- 0                           # adjustment for multi-arm trials is zero for control
arm
  delta[i,1] <- 0                       # treatment effect is zero for control arm
  mu[i] ~ dnorm(0,.0001)                # vague priors for all trial baselines

for (k in 1:na[i]) {                    # LOOP THROUGH ARMS
  r[i,k] ~ dbin(p[i,k],n[i,k])          # binomial likelihood
  logit(p[i,k]) <- mu[i] + delta[i,k]   # model for linear predictor
  rhat[i,k] <- p[i,k] * n[i,k]          # expected value of the numerators

#Deviance contribution
   dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))
     + (n[i,k]-r[i,k]) * (log(n[i,k]-r[i,k]) - log(n[i,k]-rhat[i,k]))) }

# Summed residual deviance contribution for this trial
 resdev[i] <- sum(dev[i,1:na[i]])
 for (k in 2:na[i]) {                   # LOOP THROUGH ARMS

# Trial-specific LOR distributions
    delta[i,k] ~ dnorm(md[i,k],taud[i,k])

# mean of LOR distributions (with multi-arm trial correction)
    md[i,k] <- d[t[i,k]] - d[t[i,1]] + sw[i,k]
# Precision of LOR distributions (with multi-arm trial correction)
    taud[i,k] <- tau *2*(k-1)/k

# Adjustment for multi-arm RCTs
    w[i,k] <- (delta[i,k] - d[t[i,k]] + d[t[i,1]])

# Cumulative adjustment for multi-arm trials
    sw[i,k] <- sum(w[i,1:k-1])/(k-1)
  }
 }
totresdev <- sum(resdev[])              # Total Residual Deviance
d[1]<-0                                 # treatment effect is zero for reference treatment

# Vague priors for treatment effects
for (k in 2:nt){ d[k] ~ dnorm(0,.0001) }
sd ~ dunif(0,5)                         # vague prior for between-trial SD
tau <- pow(sd,-2)                       # between-trial precision = (1/between-trial variance)



# Pairwise ORs and LORs for all possible pair-wise comparisons, if nt>2
for (c in 1:(nt-1)) {  for (k in (c+1):nt) {
or[c,k] <- exp(d[k] - d[c])
lor[c,k] <- (d[k]-d[c])
}
}

# Ranking on relative scale
for (k in 1:nt) {
rk[k] <- rank(d[],k)                    # assumes events are "bad"
best[k] <- equals(rk[k],1)              #calculate probability that treat k is best
}
}                                       # PROGRAM ENDS
```

6.6 Findings (Systematic Review)

6.6.1 Study selection

The PRISMA flow diagram (see Figure 40) presents the study selection process with the reasons for exclusion. The updated database search revealed a potential 165 studies in addition to the 44 studies identified by Maman et al (total of 209 studies). After removing duplicates (39 studies), 170 references were obtained for full text review and screening against the inclusion and exclusion criteria. This resulted in excluding a further 149 articles (total of 188 studies excluded) which resulted in 21 studies which fulfilled the eligibility criteria for the MTC analysis. These 21 studies enrolled 21,855 patients of which 18,863 patients represented allocations to treatment arms of a medicine and dose licensed within the UK.(208-227)

6.6.2 Study characteristics

The main characteristics of the studies included within the MTC analysis are documented in Table 21. Studies which met the inclusion / exclusion criteria contained the following pharmacological treatments: 15mg darifenacin; 4mg and 8 mg fesoterodine; 50mg mirabegron; 10mg modified-release oxybutynin ; 5mg and 10mg solifenacin; 4mg immediate-release tolterodine; 4mg modified-release tolterodine; 40mg immediate-release trospium; 60 mg modified-release trospium and placebo. All studies contained a minimum of two arms, where the control group could use an antimuscarinic (different drug, formulation, or dosage) or placebo. Figure 43 provides an overview of the treatment network, including the number of direct comparisons.

6.6.3 Quality assessment of the included studies

Using the Cochrane risk of bias assessment tool, none of the 21 trials included were identified at a high risk of bias for any of the domains; all studies were therefore included in the MTC analyses. The assessment for each study is described in Table 22 and Table 23 below.

Table 21: Characteristics of studies included within the mixed-treatment comparison analysis for overactive bladder syndrome

Study	Intervention	Trial design	Patient population	No randomised patients	Trial length (weeks)
Chapple 2013 (DRAGON)**(208)**	• Mirabegron 25mg • Mirabegron 50mg • Mirabegron 100mg • Mirabegron 200mg • Tolterodine ER 4mg • Placebo	Phase IIb, RCT, double-blind, double dummy, multicentre (Europe)	OAB, ≥ 18 years	928	12
Khullar 2013 (SCORPIO)**(209)**	• Mirabegron 50mg • Mirabegron 100mg • Tolterodine ER 4mg • Placebo	Phase III, RCT, double-blind, multicentre (Europe, Australia)	OAB, ≥ 18 years	1987	12
Nitti 2013 (ARIES)**(210)**	• Mirabegron 50mg • Mirabegron 100mg • Placebo	Phase III, RCT, double-blind, double dummy, multicentre (US and Canada)	OAB, ≥ 18 years	1329	12
Herschorn 2013 (CAPRICORN)**(211)**	• Mirabegron 25mg • Mirabegron 50mg • Placebo	Phase III, RCT, double-blind, double dummy, multicentre (Europe, US, Canada)	OAB, ≥ 18 years	1306	12
Yamaguchi 2014 (178-CL-048)**(212)**	• Mirabegron 50mg • Tolterodine tartrate 4mg • Placebo	RCT, double-blind, multicentre (Japan)	OAB , ≥ 20 years	1139	12
Appell 2001**(213)**	• Tolterodine IR 2mg BD • Oxybutynin ER 10mg	RCT, double-blind, multicentre (USA)	OAB	378	12
Cardozo 2004**(214)**	• Solifenacin 5mg • Solifenacin 10mg • Solifenacin 20mg	RCT, double-blind, multicentre	OAB, ≥ 18 years	907	12
Chapple 2004**(215)**	• Solifenacin 5mg • Solifenacin 10mg • Tolterodine 2mg BD • Placebo	Phase IIIa RCT, double-blind, multicentre (North America & Europe)	OAB, ≥ 18 years	1081	12

Study	Intervention	Trial design	Patient population	No randomised patients	Trial length (weeks)
Choo 2008**(228)**	• Solifenacin 5mg • Solifenacin 10mg • Tolterodine IR 2mg BD	Phase III RCT, double-blind, multicentre (Korea)	OAB, ≥ 18 years	329	12
Chu 2009**(216)**	• Solifenacin 10mg • Placebo	Phase III RCT, double-blind, multicentre (USA)	OAB, ≥ 18 years	672	12
Herschorn 2008**(217)**	• Tolterodine ER 4mg • Placebo	RCT, double-blind, multicentre (Canada, Europe)	OAB, ≥ 18 years	617	12
Herschorn 2010**(218)**	• Fesoterodine 4/8mg • Tolterodine ER 4mg • Placebo	RCT, double-blind, double dummy, multicentre (USA)	OAB, ≥ 18 years	1712	12
Homma 2003**(219)**	• Tolterodine ER 4mg • Oxybutynin 3mg TDS • Placebo	RCT, double-blind, multicentre (Japan)	OAB, ≥ 20 years	608	12
Kaplan 2011**(220)**	• Fesoterodine 4mg/8mg • Tolterodine ER 4mg • Placebo	RCT, double-blind, double dummy, multicentre (North & South America, Europe, Asia, Africa)	OAB, ≥ 18 years	2417	12
Nitti 2007**(221)**	• Fesoterodine 4mg • Fesoterodine 8mg • Placebo	RCT, double-blind, multicentre (USA)	OAB, ≥ 18 years	836	12
Staskin 2007**(222)**	• Trospium chloride 60mg • Placebo	RCT, double-blind, multicentre (USA)	Subjects with OAB	601	12
Van Kerrebroeck 2001**(223)**	• Tolterodine IR 2mg BD • Tolterodine ER 4mg • Placebo	RCT, double-blind, multicentre (Australasia, Europe, North America)	OAB, ≥ 18 years	1529	12
Yamaguchi 2007**(224)**	• Solifenacin 5mg • Solifenacin 10mg • Propiverine 20mg • Placebo	Phase III RCT, double-blind, multicentre (Japan)	OAB, ≥ 20 years	1584	12

Study	Intervention	Trial design	Patient population	No randomised patients	Trial length (weeks)
Yamguchi 2011**(225)**	• Fesoterodine 4mg • Fesoterodine 8mg • Placebo	Phase III RCT, double-blind, multicentre (Asia)	OAB, ≥ 20 years	951	12
Zinner 2004**(226)**	• Trospium 40mg • Placebo	RCT, multicentre (USA)	OAB, ≥ 18 years	512	12
Zinner 2006**(227)**	• Darifenacin 15 mg • Placebo	RCT, double-blind, multicentre	OAB, ≥ 18 years	432	12

Table 22: Risk of bias assessment tool

Study Questions	Description
Q1	Was randomisation carried out appropriately?
Q2	Was the concealment of treatment allocation adequate?
Q3	Were the groups similar at the outset of the study in terms of prognostic factors?
Q4	Were the care providers, participants and outcome assessors blind to treatment allocation?
Q5	Were there any unexpected imbalances in drop-outs between groups?
Q6	Is there any evidence to suggest that the authors measured more outcomes than they reported?
Q7	Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?

Table 23: Trials included within the analysis and their risk of bias assessment

	Study questions						
	Q1	Q2	Q3	Q4	Q5	Q6	Q7
Chapple 2013 (DRAGON)(208)	Yes	Yes	Yes	Yes	No	No	Yes
Khullar 2013 (SCORPIO)(209)	Yes	Yes	Yes	Yes	No	No	Yes
Nitti 2013 (ARIES)(210)	Yes	Yes	Yes	Yes	No	No	Yes
Herschorn 2013 (CAPRICORN)(211)	Yes	Yes	Yes	Yes	No	No	Yes
Yamaguchi 2014 (175-CL-048)(212)	Yes	Yes	Yes	Yes	No	No	Yes
Appell 2001(213)	Unclear	Unclear	Yes	Yes	No	No	Yes
Cardozo 2004(214)	Unclear	Unclear	Yes	Unclear	No	No	Yes
Chapple 2004(215)	Unclear	Unclear	Yes	Unclear	No	Yes	Yes
Choo 2008(228)	Unclear	Unclear	Yes	Unclear	No	No	Yes
Chu 2009(216)	Yes	Yes	Yes	Yes	No	No	Yes
Herschorn 2008(217)	Unclear	Unclear	Yes	Unclear	No	No	Yes
Herschorn 2010(218)	Yes	Yes	Yes	Unclear	No	No	Yes
Homma 2003(219)	Yes	Unclear	Yes	Yes	No	No	Yes
Kaplan 2011(220)	Yes	Unclear	Yes	Yes	No	No	Yes
Nitti 2007(221)	Yes	Unclear	Yes	Unclear	No	No	Yes
Staskin 2007(222)	Unclear	Unclear	Yes	Unclear	No	No	Yes
Van Kerrebroeck 2001(223)	Yes	Unclear	Yes	Unclear	No	No	Yes
Yamaguchi 2007(224)	Unclear	Unclear	Yes	Unclear	No	No	Yes
Yamaguchi 2011(225)	Unclear	Unclear	Yes	Unclear	No	No	Yes
Zinner 2004(226)	Yes	Unclear	Yes	Unclear	No	No	Yes
Zinner 2006(227)	Yes	Unclear	Yes	Unclear	No	No	Yes

Figure 43: Antimuscarinics included within the network meta-analysis / mixed-treatment comparison.

Each treatment represents a node within the star-shaped network. The lines between the nodes represent direct comparative data, where the number along the line indicates the number of studies for that particular link within the network. Each triangle represents a loop of direct comparative data which allows mixed treatment comparison.

6.7 Findings (Bayesian MTC)

6.7.1 Efficacy

6.7.1.1 Micturition per 24 hours

The MTC analysis on micturition frequency was based on all 21 studies (18,863 patients).(208-228) The mean (SD) baseline micturition frequency per 24 hours was 11.6 (±2.4). The median placebo-corrected difference from baseline to week 12 was statistically significantly different for all treatments except darifenacin 15mg (see Figure 44). 10mg modified-release oxybutynin demonstrated the greatest reduction in micturition frequency with a posterior median of the mean difference of -1.38 (95% CrI -1.79 to -0.97) with 10mg solifenacin (-1.31; 95% CrI -1.55 to -1.06) demonstrating the second greatest reduction. For the ranking analysis, 10mg solifenacin was assigned as the best agent (see Table 24) due to greater confidence in its estimate from the MTC analysis over 10mg modified-release oxybutynin, as evinced by narrower confidence intervals.

6.7.1.2 Urinary urgency incontinence (UUI) per 24 hours

The MTC analysis on UUI episodes was based on 17 studies (15,502 patients).(208-215, 217, 218, 220-222, 224, 225, 227) The mean (SD) baseline UUI per 24 hours was 3.7 (±2.1). The median placebo-corrected difference from baseline to week 12 was statistically significantly different for all treatments except 4mg immediate-release tolterodine, 4mg fesoterodine, 40mg immediate-release trospium, and 15mg darifenacin (see Figure 45). 10mg solifenacin demonstrated the greatest reduction in number of UUI episodes with a posterior median of the mean difference of -0.91 (95% CrI -1.25 to -0.56) with solifenacin 5mg (-0.67; 95% CrI -1.02 to -0.32) and 10mg modified-release oxybutynin (-0.66, 95% CrI -1.20 to -0.14) demonstrating joint second greatest reduction. For the ranking analysis, 10mg solifenacin was assigned as the best agent (see Table 25).

6.7.2 Adverse events

6.7.2.1 Dry mouth

The MTC analysis on incidence of dry mouth was based on all 21 studies (18,863 patients). (208-228) 50mg Mirabegron had an incidence of dry mouth similar to placebo (OR 0.75; 95% CrI 0.50 to 1.16). All antimuscarinics were associated with a significantly higher risk of dry mouth compared with placebo and mirabegron (see Figure 46). For the ranking analysis, 50mg mirabegron was assigned as the best agent (see Table 26).

6.7.2.2 Constipation

The MTC analysis on incidence of constipation was based on 19 studies (17,454 patients).(208, 209, 212-228) 50mg Mirabegron, 10mg modified-release oxybutynin, and 4mg Fesoterodine had a lower (not reaching significance) incidence of constipation compared with placebo. All other antimuscarinics were associated with a significantly higher risk of constipation (see Figure 47). For the ranking analysis, Fesoterodine 4mg was assigned as the best agent (see Table 27).

6.7.3 Rank

A summary of the ranking analysis for each of the four endpoints is presented in Table 28. Detailed distribution of probabilities (Rank-o-gram) for each treatment being ranked at each of the possible 12 positions is depicted in Figure 48 using the assumption that each endpoint is equally weighted to each other. Solifenacin and oxybutynin were among the most efficacious treatments, whilst mirabegron was the best tolerated.

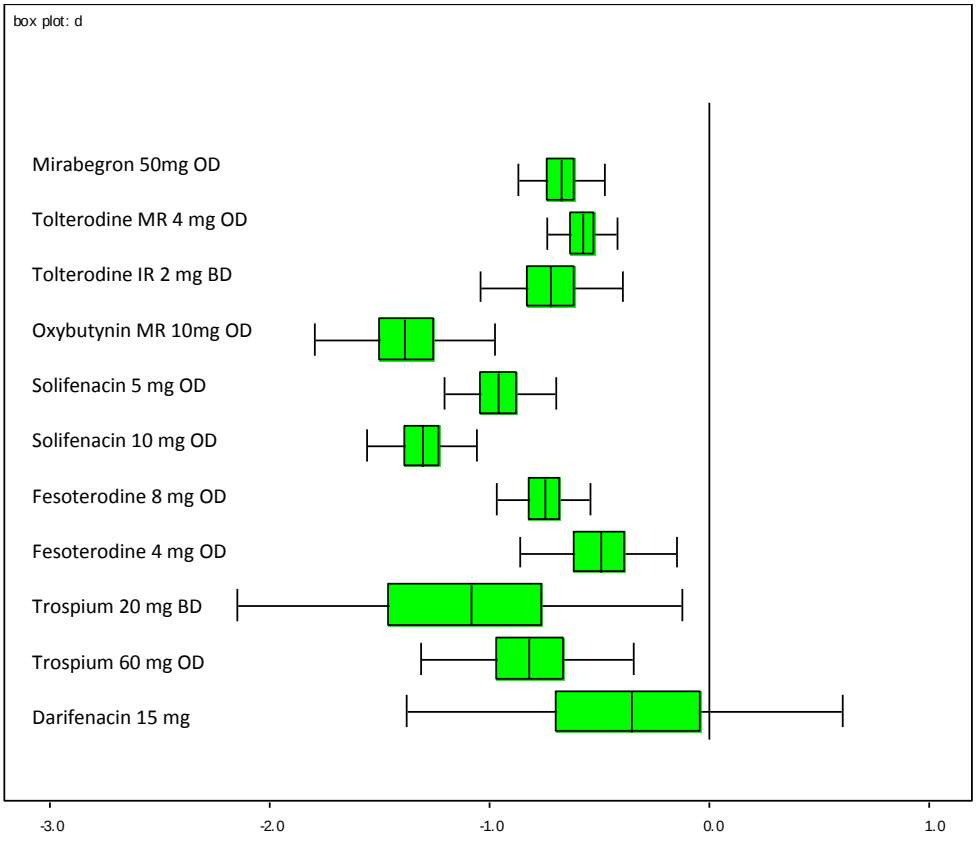


Figure 44: Mean change from baseline (micturition frequency per 24 hours) versus placebo

Box plots of the placebo-corrected mean difference of pharmacological treatments for OAB as estimated by the Bayesian MTC analysis regarding efficacy (micturition frequency per 24 hours). The individual plots represent the 95% credible interval (horizontal black line either side of the green box) and the interquartile range (green box, representing where 25% to 75% of the data lies); the vertical black line reflects the median placebo-corrected difference (MD). A larger MD (in the negative direction) is indicative of a greater treatment effect, relative to placebo. Where the 95% credible interval crosses the line of unity, this is indicative of a non-statistically significant difference to placebo..

Table 24: Mean difference and rank order of the effect of treatment on micturition frequency per 24 hours

Drug (dose)	Mean difference	95% Credible Interval	Rank
Mirabegron 50mg OD	-0.674	-0.865, -0.478	7
Tolterodine MR 4 mg OD	-0.571	-0.739, -0.416	9
Tolterodine IR 2 mg BD	-0.721	-1.040, -0.397	7
Oxybutynin MR 10 mg OD	-1.381	-1.792, -0.974	2
Solifenacin 5 mg OD	-0.960	-1.205, -0.698	4
Solifenacin 10 mg OD	-1.305	-1.552, -1.056	1
Fesoterodine 8 mg OD	-0.747	-0.967, -0.538	6
Fesoterodine 4 mg OD	-0.494	-0.860, -0.145	10
Trospium 20 mg BD	-1.080	-2.146, -0.122	3
Trospium 60 mg OD	-0.816	-1.312, -0.345	5
Darifenacin 15 mg	-0.355	-1.378, 0.608	11

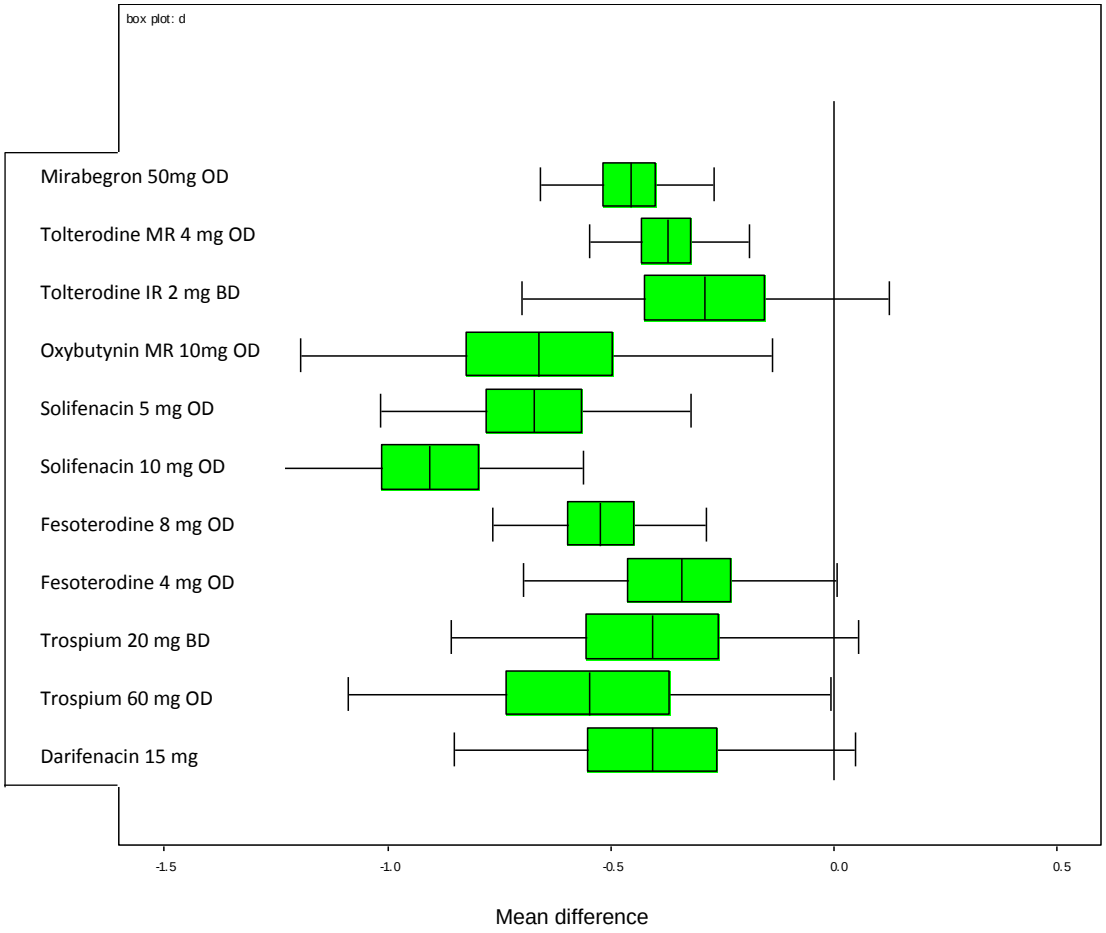


Figure 45: Mean change from baseline (number of UUI episodes per 24 hours) versus placebo

Box plots of the placebo-corrected mean difference of pharmacological treatments for OAB as estimated by the Bayesian MTC analysis regarding efficacy (UUI episodes per 24 hours). The individual plots represent the 95% credible interval (horizontal black line either side of the green box) and the interquartile range (green box, representing where 25% to 75% of the data lies); the vertical black line reflects the median placebo-corrected difference (MD). A larger MD (in the negative direction) is indicative of a greater treatment effect, relative to placebo. Where the 95% credible interval crosses the line of unity, this is indicative of a non-statistically significant difference to placebo.

Table 25: Mean difference and rank order of the effect of treatment on UUI frequency per 24 hours

Drug (dose)	Mean difference	95% Credible Interval	Rank
Mirabegron 50mg OD	-0.456	-0.658, -0.268	6
Tolterodine MR 4 mg OD	-0.374	-0.549, -0.190	7
Tolterodine IR 2 mg BD	-0.291	-0.701, 0.122	11
Oxybutynin MR 10 mg OD	-0.660	-1.196, -0.137	2
Solifenacin 5 mg OD	-0.672	-1.017, -0.322	2
Solifenacin 10 mg OD	-0.907	-1.246, -0.563	1
Fesoterodine 8 mg OD	-0.522	-0.764, -0.286	4
Fesoterodine 4 mg OD	-0.343	-0.697, 0.005	10
Trospium 20 mg BD	-0.405	-0.857, 0.053	7
Trospium 60 mg OD	-0.548	-1.089, -0.008	4
Darifenacin 15 mg	-0.406	-0.849, 0.050	7

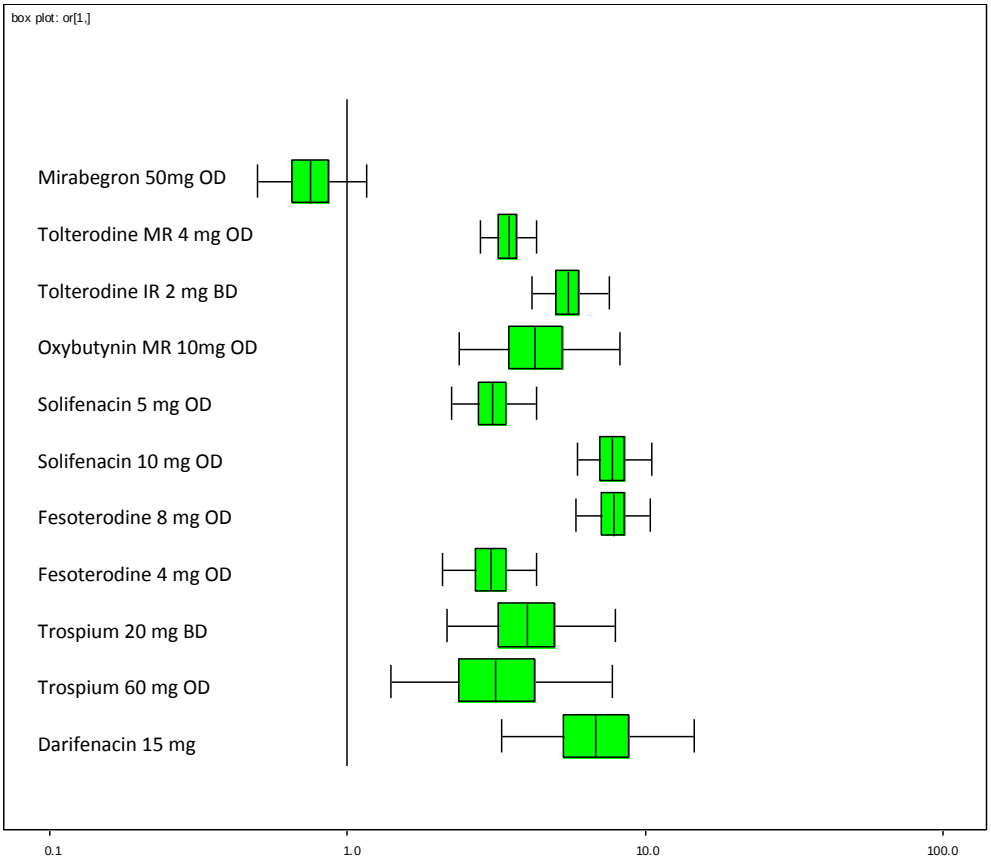


Figure 46: Odds ratio (median) incidence of dry mouth relate to placebo

Box plots of the placebo-corrected odds ratio (OR) of pharmacological treatments for OAB as estimated by the Bayesian MTC analysis regarding tolerability (dry mouth). The individual plots represent the 95% credible interval (horizontal black line either side of the green box) and the interquartile range (green box, representing where 25% to 75% of the data lies); the vertical black line reflects the median placebo-corrected difference (MD). A larger OR (in the positive direction) is indicative of a greater adverse effect, relative to placebo. Where the 95% credible interval crosses the line of unity, this is indicative of a non-statistically significant difference to placebo.

Table 26: Odds ratio and rank order of the effect of treatment on dry mouth

Drug (dose)	Odds Ratio	95% Credible Interval	Rank
Mirabegron 50mg OD	0.7528	0.4982, 1.157	1
Tolterodine MR 4 mg OD	3.479	2.793, 4.337	5
Tolterodine IR 2 mg BD	5.545	4.153, 7.588	8
Oxybutynin MR 10 mg OD	4.283	2.384, 8.221	7
Solifenacin 5 mg OD	3.086	2.233, 4.318	3
Solifenacin 10 mg OD	7.778	5.896, 10.550	10
Fesoterodine 8 mg OD	7.896	5.852, 10.420	10
Fesoterodine 4 mg OD	3.043	2.084, 4.323	3
Trospium 20 mg BD	4.048	2.156, 7.986	7
Trospium 60 mg OD	3.162	1.407, 7.724	4
Darifenacin 15 mg	6.809	3.312, 14.610	9

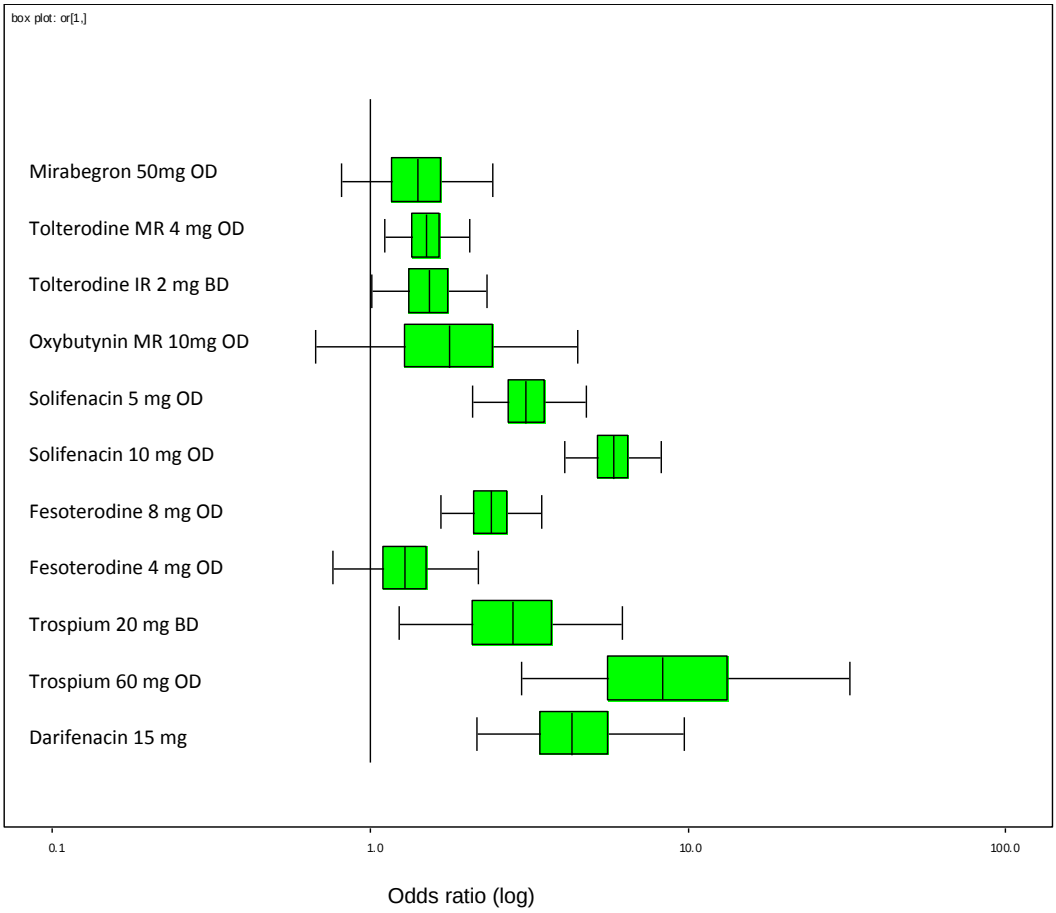


Figure 47: Odds ratio (median) incidence of constipation relative to placebo

Box plots of the placebo-corrected odds ratio (OR) of pharmacological treatments for OAB as estimated by the Bayesian MTC analysis regarding tolerability (constipation). The individual plots represent the 95% credible interval (horizontal black line either side of the green box) and the interquartile range (green box, representing where 25% to 75% of the data lies); the vertical black line reflects the median placebo-corrected difference (MD). A larger OR (in the positive direction) is indicative of a greater adverse effect, relative to placebo. Where the 95% credible interval crosses the line of unity, this is indicative of a non-statistically significant difference to placebo.

Table 27: Odds ratio and rank order of the effect of treatment on constipation

Drug (dose)	Mean difference	95% Credible Interval	Rank
Mirabegron 50mg OD	1.412	0.812, 2.422	4
Tolterodine MR 4 mg OD	1.496	1.106, 2.051	4
Tolterodine IR 2 mg BD	1.534	1.005, 2.332	4
Oxybutynin MR 10 mg OD	1.781	0.672, 4.492	5
Solifenacin 5 mg OD	3.093	2.102, 4.784	8
Solifenacin 10 mg OD	5.826	4.098, 8.260	10
Fesoterodine 8 mg OD	2.394	1.661, 3.453	6
Fesoterodine 4 mg OD	1.280	0.760, 2.187	1
Trospium 20 mg BD	2.819	1.234, 6.222	7
Trospium 60 mg OD	8.347	3.000, 32.370	11
Darifenacin 15 mg	4.295	2.169, 9.677	9

Table 28: Trade-off analysis – summary rank of treatment (efficacy and tolerability) for the management of overactive bladder syndrome

	Micturition / 24hrs	UUI / 24hrs	Dry mouth	Constipation
Mirabegron 50mg OD	7	6	1	2
Tolterodine MR 4mg OD	9	7	5	2
Tolterodine IR 2mg BD	7	11	8	2
Oxybutynin MR 10mg OD	2	2	6	5
Solifenacin 5mg OD	4	2	2	8
Solifenacin 10mg OD	1	1	10	10
Fesoterodine 8mg OD	6	4	10	6
Fesoterodine 4mg OD	10	10	2	1
Trospium 20mg BD	3	7	6	7
Trospium 60mg OD	5	4	4	11
Darifenacin 15mg OD	11	7	9	9

Table 29: Trade-off analysis – summary rank of treatment (clinical vs. cost)

	Trade-Off (clinical) Rank	Cost per 28-days excl. VAT (BNF)(229)	Cost Rank	Revised Rank (Top 3)
Mirabegron 50mg OD	3	27.06	4	--
Tolterodine MR 4mg OD	6	25.78	4	--
Tolterodine IR 2mg BD	8	2.65	2	2
Oxybutynin MR 10mg OD	2	25.71	4	--
Solifenacin 5mg OD	1	25.78	4	3
Solifenacin 10mg OD	4	33.52	5	3
Fesoterodine 8mg OD	7	25.78	4	--
Fesoterodine 4mg OD	5	25.78	4	--
Trospium 20mg BD	5	23.77	3	--
Trospium 60mg OD	5	23.05	3	--
Darifenacin 15mg OD	9	25.48	4	--
Oxybutynin IR 5mg BD	(2)	2.47	1	1

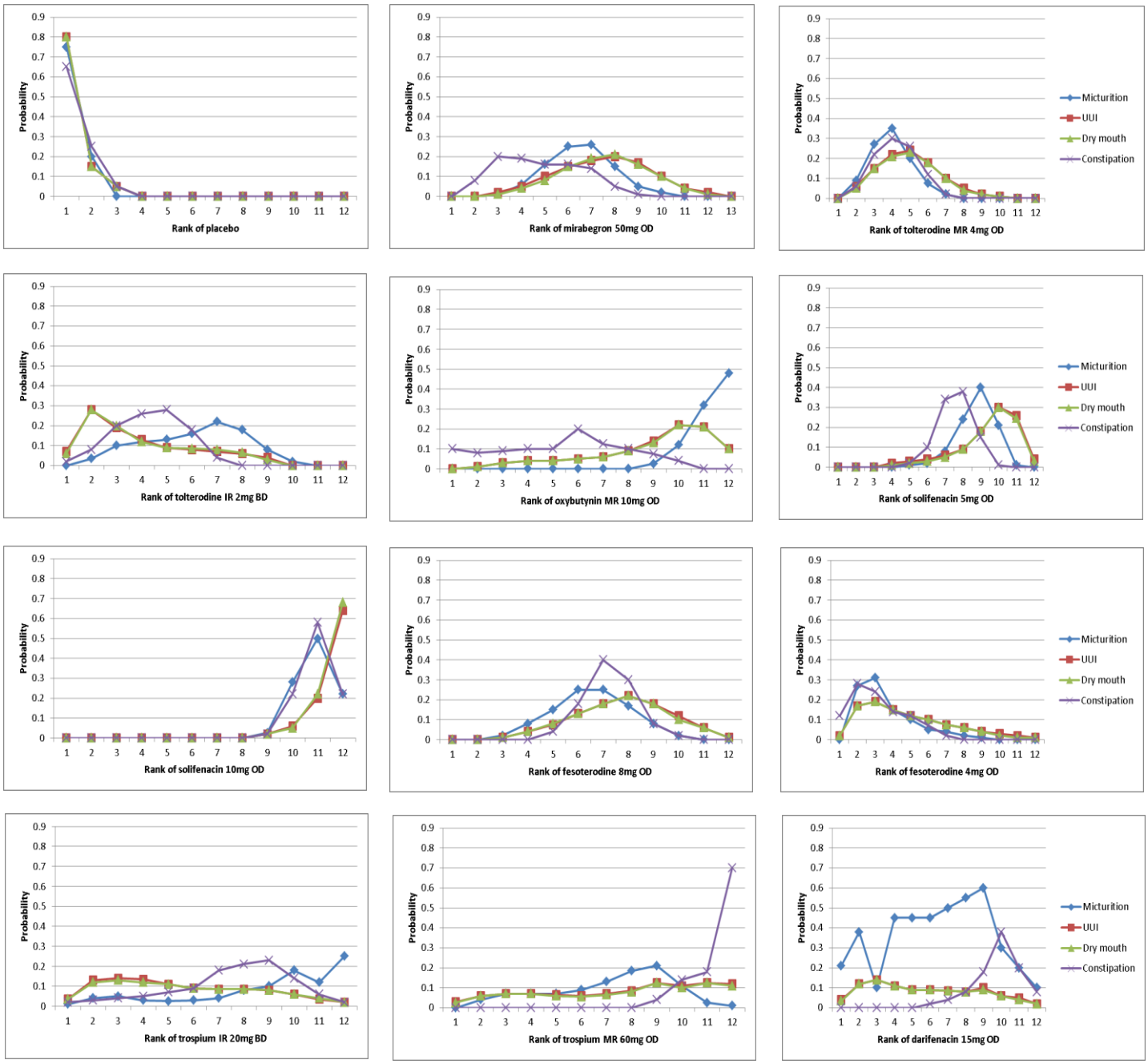


Figure 48: Trade-off analysis – rank-o-gram for each of the efficacy and tolerability endpoints

Ranking indicates the probability to be the best treatment, the second best, the third best, and so on, among the 11 pharmacological agents (and placebo) for the management of overactive bladder. A higher probability at a rank of 12 is the most favourable for the efficacy endpoints (micturition and UUI, whilst a higher probability at a rank of 1 is the most favourable for the tolerability endpoints (dry mouth and constipation). For example, placebo demonstrated the number 1 rank for micturition and UUI indicating that it was the least effective for efficacy, however, placebo also had a rank of 1 for constipation indicating that it was the best tolerated for this safety endpoint.

6.7.4 Quality assessment of the MTC model

For all four outcomes, the model showed reasonable goodness of fit to the data (number of data points) as determined by the posterior mean of the residual deviance [micturition = 52.7 (55), p = 0.564; UUI = 39.9 (45), p = 0.689; dry mouth = 45.5 (54), p = 0.788; constipation = 40.6 (50), p = 0.826].

No evidence of significant inconsistency (Bucher's test) was detected between directly observed and inferred treatment effects within the loop identified in Figure 43 for any of the four endpoints. A visual assessment of the trace plots (history) and time series (density) plots also did not reveal cause for concern regarding inconsistency.

6.8 Summary

Based on the data from the 21 studies included within the analyses, the use of anticholinergics or mirabegron for the management of OAB reduces symptoms by a potentially clinically insignificant value. For example, for the micturition frequency per 24 hours endpoint, the greatest reduction was seen with oxybutynin (placebo corrected reduction of -1.38 against a mean baseline of 11.6 per 24 hours). Nonetheless, this MTC and trade-off analysis suggests that 10mg solifenacin once daily is the most *effective* treatment for symptoms of OAB (reduction in micturition frequency and number of UUI episodes per 24 hours) compared with other agents available within the UK for this indication. However, 10mg solifenacin, along with 8mg Fesoterodine, were among the treatments with the highest incidence of dry mouth, and along with 60mg modified-release trospium, were among the treatments with the highest incidence of constipation. The agent with the most favourable *tolerability* profile was 50mg mirabegron, which was shown to be associated with a risk of dry mouth similar to placebo and significantly lower compared to all antimuscarinic drugs. This is unsurprising given its mechanism of action is different i.e. not acting on the anticholinergic pathway. The MTC also suggested that the risk of constipation was lower with 50mg mirabegron than most antimuscarinics (excluding tolterodine).

This MTC analysis confirms the finding of previous systematic review and meta-analyses conducted by Chapple et al.,(203) and Maman et al.(205) The merits of this review over the previous two published in recent years are two-fold. First, it utilised a stricter inclusion and quality assessment criteria to increase the robustness of the finding of the MTC analysis; second, a quantitative trade-off assessment of treatments for each of the four endpoints investigated.

The results of the present review differ from those reported by Buser et al.(204) specifically that solifenacin was estimated as being the most efficacious treatment in preference to trospium, oxybutynin and fesoterodine. The key reason for the differences is that Buser et al. included many treatments at doses which are unconventional and unlicensed. The authors note that this was done purely for the sake of

completeness with the corresponding results being incompletely understood and irrelevant to clinical practice. Irrespective of these differences in inclusion criteria, both meta-analyses suggest that there are few statistically significant important differences in efficacy between the agents, highlighting that 10 mg solifenacin ranked first with regards to efficacy (micturition frequency and UUI episodes) among pharmacological treatments included within the present analysis.

6.8.1 Consideration of the trade-off analysis (oxybutynin modified-release vs. immediate release)

Although the results of the analysis reports oxybutynin modified-release and solifenacin as being the agents possessing the best clinical outcomes (micturitions per 24 hours and UUI per 24 hours), with the exception of darifenacin, this difference relates to at best, one less micturition episode per 24 hours or one-half less UUI episode per 24 hours. On this basis, given the high cost of these two agents (currently marketed under a protected patent), it is unlikely that this difference constitutes a clinically meaningful difference to prescribe these in preference to the generic agents, which are equally efficacious.

The authors of the Cochrane Collaboration review (2012) concluded the following:

"Where the prescribing choice is between oral immediate-release oxybutynin or tolterodine, tolterodine might be preferred for reduced risk of dry mouth. If extended-release preparations of oxybutynin or tolterodine are available, these might be preferred to immediate-release preparations because there is less risk of dry mouth. Between solifenacin and immediate-release tolterodine, solifenacin might be preferred for better efficacy and less risk of dry mouth. Solifenacin 5 mg once daily is the usual starting dose, this could be increased to 10 mg once daily for better efficacy but with increased risk of dry mouth. Between fesoterodine and extended-release tolterodine, fesoterodine might be preferred for superior efficacy but has higher risk of withdrawal due to adverse events and higher risk of dry mouth. There is little or no evidence available about quality of life, costs, or long-

term outcome in these studies. There were insufficient data from trials of other anticholinergic drugs to draw any conclusions." (230)

Although oxybutynin immediate-release was not included within the current analysis as no studies met the inclusion criteria, it is worth highlighting the findings of the following reports which did not meet the inclusion criteria:

- Minassiuan et al (2007) conducted a prospective, randomised, 12-week, *open-label* study.(231) The primary endpoint was number of number of micturitions per 24 hours. Although an *a priori* sample size of 60 subjects per group was required the authors only manage to recruit 72 women due to recruitment difficulties. For this reason, the study was terminated prematurely; however, an analysis of the subjects recruited demonstrated no difference between the two groups. Although these results are not adequately powered to definitely answer the research question, it is consistent with the pharmacology that immediate-release preparation will not be therapeutically inferior to the modified-release preparation.
- Barkin et al (2004) conducted a multicentre, double-blind *6-week* study.(232) Efficacy endpoints evaluated included micturition and UUI episodes. Of the 125 subjects randomised, the intention-to-treat population comprised of 94 (75%), with tolerability assessment in all. Both treatments produced equivalent reductions in micturition and UUI episodes ($p<0.001$ vs. baseline), although more subjects in the immediate-release group reported adverse effects with dry mouth being the most common.
- Versi et al (2000) conducted a multicentre, double-blind study of 226 subjects to evaluate the efficacy and safety of the modified-release preparation compared with the immediate-release preparation, with a particular emphasis on dry mouth.(233) Reductions in UUI episodes from baseline to end of treatment were 18.6 to 2.9 per week (83% decrease) and 19.8 to 4.4 per week (76% decrease) in the modified-release and immediate-release groups, respectively ($p=0.36$). At equal doses, comparable proportions of patients in both groups reported the

absence of urge incontinence (p=0.85). The incidence of dry mouth in the study was shown to increase with dose in both groups, however, there was no significant difference between the two groups (47.7% and 59.1%, respectively (p=0.9) although time to first report of dry mouth was higher in the immediate-release group.

- Anderson et al (1999) conducted a multicentre, randomised, double-blind, parallel study.(234) The primary endpoint was UUI episodes. A total of 97 women and 8 men were enrolled. The number of weekly UUI episodes decreased from 27.4 to 4.8 in the group taking the modified-release preparation compared with a decrease from 23.4 to 3.1 in the group taking the immediate-release preparation (p=0.56) demonstrating no difference between the two formulations. Total incontinence episodes decreased from 29.3 to 6.0 and from 26.3 to 3.8, respectively (p=0.6), also showing no difference between the two formulations. Continence was achieved in 41% of subjects in the modified-release preparation group compared with 40% of subjects in the immediate-release preparation group (p=0.9). Regarding tolerability, dry mouth was reported in 68% and 87% of subjects taking modified-release and immediate-release oxybutynin, respectively (p=0.04), and moderate-severe dry mouth in 25% and 46%, respectively (p=0.03).

Although studies investigating oxybutynin immediate-release did not meet the inclusion criteria, the conclusions from the above qualitative assessment suggests that the conclusions drawn from oxybutynin modified-release within the trade-off analysis can be interchangeable with oxybutynin immediate-release.

6.8.2 Consideration of the trade-off analysis vs. cost (Primary Care)

The trade-off analysis (presented in Table 28) which considered clinical endpoints was updated with cost as an additional parameter. From this, a revised top 3 rank was established (see Table 29). With consideration to efficacy, tolerability and cost, the analysis suggests that oxybutynin immediate-release should be the first-line option, with

the modified-release preparation strictly reserved for patients who have demonstrated improvements in symptoms but report issues with compliance. Where oxybutynin is not appropriate, due to concerns regarding frailty, tolterodine may be considered an alternative first-line option. Although solifenacin performed the best on the trade-off analysis between efficacy and tolerability, due to its high cost relative to generic oxybutynin / tolterodine, the revised rank places this as second / third-line.

6.8.3 Strengths and weaknesses

The methods applied in this paper allow steps beyond conventional meta-analysis. The approach incorporates all available information from clinical trials while fully maintaining randomisation. A comparison of the various treatments available and a rank ordering is estimated. Finally, the quantitative trade-off assessment to produce a ranking for the efficacy and safety endpoints provides an evidence base for formal decision-analytic models.

The use of stricter inclusion and quality assessment criteria has been utilised to overcome the limitation of the previously published reviews which included studies with poor reporting quality necessitating the authors to impute missing values (such as standard deviation).

The safety dataset utilised for the MTC analysis was based on averaged data from the published manuscripts. It is therefore unclear how predictive these mean values are for the individual patients. Further, none of the studies reported whether, or how many, patients had two or more adverse events. Therefore, it was assumed that the occurrence of an adverse event was independent of the presence of another adverse event, although this situation might be more complex in clinical practice. Almost all studies published are fixed-dose trials; the few studies that did not clearly specify the dose regarding adverse events were excluded. This limitation is important, since in clinical practice the dose of a drug is often titrated, and flexible dose studies tend to report fewer adverse events.

One possible criticism of the current analysis is that the MTC analysis for the safety outcomes focused only on dry mouth and constipation, adverse events known to be associated with antimuscarinics and not associated with mirabegron. However, on the basis of pooled safety analyses for mirabegron showing no statistically significant differences compared with placebo for adverse events reported (except nasopharyngitis) the opinion put forward is that this limitation is upheld.

The efficacy outcomes considered in this MTC were limited to micturition frequency and UUI episodes frequency, and did not include outcomes analysed in other reviews such as urgency, volume voided per micturition, or dry rate. Although urgency is an important symptom in OAB, measurement of this symptom is difficult due to the subjective nature of urgency. As several instruments can be used to measure this symptom it makes comparisons between studies challenging. The volume voided per micturition was not included as it was considered a surrogate outcome measure of less clinical relevance. Other efficacy outcomes such as number of patients with resolution of incontinence, mean change from baseline in daytime micturition frequency, night-time micturition frequency, and urgency-driven micturition are also relevant but less frequently reported, and therefore they were not considered in this MTC. A further criticism may be that this study takes no account of adjunctive therapies such as bladder training or fluid management – these may have influenced the outcome of both the studies included and the data interpreted herein although each study would have balanced this as part of the patient characteristics.

Despite the focus on objective data, the difference in the period of time on the open market for individual agents may have affected the volume of data for different agents. This may have impacted upon the study outcome and continues to be a limitation of NMA in general for the analysis of newer vs. older agents.

Lastly, the baseline for this analysis was the systematic review reported by Maman et al in which studies published before 2000 were

not included to ensure greater homogeneity between study populations. This methodology was considered sound as pre-2000 more patients were treatment naive and the definition of OAB was introduced in 2001. The drawback of this approach is that this analysis omits data on older treatments, specifically oxybutynin immediate-release which is used routinely in practice.

6.8.4 Implications for clinical practice

This review is focused on the most clinically relevant factors when considering treatment for symptoms of OAB as validated by a Consultant within this field. Although it is widely accepted that the best model of providing answers with the highest clinical relevance would be from a large scale multiple comparison head-to-head trial, MTC analysis are increasingly used to provide guidance in areas where a multitude of therapeutics options are available. An important implication of the findings is that solifenacin is the highest ranking treatment with regards to the trade-off analysis of clinical (efficacy and tolerability) parameters; however, when the absolute differences are considered in relation to cost of treatment, this revised analysis places solifenacin as a second-line option behind oxybutynin.

6.8.5 Conclusion

This analysis found that all treatments for OAB were associated with a reduction in symptoms; however, not all were statistically significant. Solifenacin 10mg was found to possess the best efficacy profile, whilst mirabegron possessed the most favourable tolerability profile. The Bayesian MTC methodology is a powerful tool for evidence-based health care evaluation, and when supplemented with a quantitative trade-off assessment, it has the ability to provide a reliable hierarchy of treatment options. The strength of the findings lies in the model being based on a robust dataset following principles from the PRISMA statement and the Cochrane Collaboration.

6.9 Peer review

Prior to performing the analysis in WinBUGS, the code was forwarded to Dr Nicky Welton (Reader in Statistical and Health Economic

Modelling; School of Social and Community Medicine, University of Bristol) for validation. Advice was provided on tidying up the lines of coding relating to inclusion of studies with three-arms as well as the performance of a number of further analyses: checking model fit / consistency; obtaining ranking probabilities / rank-o-gram; and including a table of treatment effect estimates. These were all performed.

The initial protocol for this project as well as the draft manuscript was forwarded to Dr Wood (Consultant, Urology) for comment and clinical support. All advise and corrections provided were implemented. Beyond this, a draft manuscript in article formal was forward to Dr Welton for comment, from which all questions and comments were addressed. The result of the methodological and clinical collaboration was to improve the overall quality of the analysis and its conclusions.

6.10 Translation of outcomes to current practice

6.10.1 Trade-off analysis

This assessment found that all treatments for OAB were associated with a similar level of efficacy and safety, with the exception of darifenacin. The trade-off analysis, with consideration to the efficacy and tolerability parameters (presented as rank-o-grams), identified solifenacin as being the most favourable agent; however, the clinical significance between the ranks was minimal due to overlapping confidence intervals. When the trade-off assessment considered cost as an additional parameter, it was agreed with clinical consultation that oxybutynin and / or tolterodine immediate-release should be recommended as first-line options.

It should be noted that the patent protection for solifenacin will cease in 2018/19 after which point it would be expected that generic (non-proprietary) products will be commercially available at a fraction of the existing cost. If the UK market price for solifenacin was to follow the same trend as seen for losartan, the price for one-month treatment of solifenacin would reduce from the current

price of £33.52 to circa £1.20 (excluding VAT) within 6 month of patent loss.

In accordance with the Health & Social Care Information Centre Prescription Cost Analysis [England 2014 data], a total of 72,070 thousand issues of solifenacin 10mg and 5 mg tablets were made in 2014 resulting in a total spend of £73,360,000. This particular medicine represents the highest value of in-class spend at 53% of the £137,275,000 total (BNF category 7.4.2). At a reduction in value in line with the proposal above (i.e. non-proprietary price of 5% that of the branded Vesicare) at current issue level, spend would reduce to £3,668,000 resulting in a saving in the region of £70m. Until such time that a generic preparation of solifenacin is available, the use of oxybutynin or tolterodine immediate-release would represent more cost-effective first-line therapies on the basis of a simple cost-minimisation principle. Refer to sections 6.10.2 and section 6.11 for current changes in practice.

6.10.2 Implications of research

The findings from this area of research were presented locally with the Area Prescribing Committee and associated Commissioners. A local guideline was written to influence and support prescribing (see section 6.10.3; NCL Guideline: Pharmacological management of OAB).

6.10.3 NCL Guideline: Pharmacological management of OAB

North Central London
Joint Formulary Committee

Pharmacological management of Overactive Bladder (OAB) Syndrome in Primary Care

Disclaimer

This guideline is registered at North Central London (NCL) Joint Formulary Committee (JFC) and is intended solely for use by healthcare professionals to aid the treatment of patients within NCL. However, clinical guidelines are for guidance only, their interpretation and application remain the responsibility of the individual clinician. If in doubt, contact a senior colleague or expert. Clinicians are advised to refer to the manufacturer's current prescribing information before treating individual patients.

The authors and NCL JFC accept no liability for use of this information from this beyond its intended use.

While we have tried to compile accurate information in this guideline, and to keep it updated in a timely manner, we cannot guarantee that it is fully complete and correct at all times. If you identify information within this guideline that is inaccurate, please report this to the admin@ncl-jfc.org.uk. If a patient is harmed as a consequence of following this guideline, please complete a local incident report and inform admin@ncl-jfc.org.uk.

This guideline is not be to used or reproduced for commercial or marketing purposes.

NCL JFC is funded by and provides advice to Acute Trusts and Clinical Commissioning Groups in NCL.

Document control

Date	Version	Amendments
19/05/2016	0.4	Updates following May 2015 JFC meeting
26/05/2016	0.5	Updates following May 2016 JFC meeting
03/06/2016	1.0	Published version
26/10/2016	1.1	Updated: Purpose amended to specific "pharmacological management" Drug selection principles amended to specify when solifenacin + mirabegron combination considered Summary of treatment recommendations updated

Document management

Groups / Individuals who have overseen the development of this guidance:	Pritesh Bodalia, JFC
Groups which were consulted and have given approval:	UCLH & RNOH Urology, Urogynaecology and Care of the Elderly teams
Version number:	1.1
Available on:	NCL JFC website
Disseminated to:	NCL CCGs
Equality impact assessment:	Low
NCL Joint Formulary Committee Approval date:	26 October 2016 (Chair's Action)
Review date:	May 2019

Contents

1. Target audience

Primary care GPs and practice nurses.

2. Purpose

To provide a treatment algorithm for the pharmacological management of Overactive Bladder Syndrome in Primary Care.

The guidance is based on best available evidence, incorporating recommendations from NICE Clinical Guideline 171 (Urinary incontinence, 2013), NICE Technology Appraisal 290 (Mirabegron, 2013). The recommendations follow an independent review of the literature (level 1, grade A evidence) by the NCL JFC and consultation with key urology stakeholders.

3. General treatment principles

- When offering anti-cholinergic drugs to treat OAB consider co-existing conditions (for example, poor bladder emptying), use of existing medication affecting the total anti-cholinergic load and risk of adverse effects.
- The exclusion of other pathologies including stones, infection or malignancy (where appropriate) is important. Refer to Section 3.1 for red flag symptoms requiring referral to secondary care
- The use of bladder diaries to assess symptoms is recommended
- Before OAB drug treatment starts, discuss with patients:
  - the likelihood of success and associated common adverse effects, and
  - the frequency and route of administration, and
  - that some adverse effects such as dry mouth and constipation may indicate that treatment is starting to have an effect, and
  - that they may not see the full benefits until they have been taking the treatment for 4 weeks
  - fluid and lifestyle advice (including caffeine and fluid reduction)
- Prescribe the lowest recommended dose when starting a new OAB drug treatment
- If OAB drug treatment is effective and well-tolerated, do not change the dose or drug

3.1. Red flag symptoms requiring referral to Secondary Care

- Visible haematuria
- Recurrent or persistent UTI associated with haematuria in women aged ≥ 40 years
- Microscopic haematuria in women aged ≥ 50 years
- Suspected malignant mass arising from the urinary tract
- Abnormal DSE or PSA
- Family history of bladder cancer
- Loss of weight
- Bone pain
- Persistent bladder or urethral pain
- Clinically benign pelvic mass

- Faecal incontinence
- Suspected neurological disease
- Voiding difficulty
- Suspected or confirmed urogenital fistulae
- Previous continence / pelvic cancer surgery
- Previous pelvic radiation therapy
- Suspected or confirmed acute kidney injury

4. Drug selection principles

- Offer oxybutynin immediate-release (**not** if frail or elderly) or tolterodine immediate-release first to patients with OAB or mixed UI who have good performance status
- If oxybutynin immediate-release is not effective or well tolerated, offer tolterodine immediate-release
- **Do not** offer oxybutynin to patients with frailty (due to potential impact on cognitive function based on crossing of the blood-brain barrier); offer tolterodine immediate-release as the first-line agent or solifenacin if an anti-cholinergic is indicated
- Review treatment after 4 weeks (refer to section 5)
- If immediate release anti-cholinergic treatment(s) for OAB or mixed UI are not effective or well tolerated, offer solifenacin
- If solifenacin alone is not effective despite dose optimisation, and the patient is unsuitable for invasive procedures, consider treatment with a combination of solifenacin and mirabegron. This combination has been shown to be effective at improving mean voided volume, micturition frequency and urgency.
- Offer mirabegron, as an alternative, if anti-cholinergics are contra-indicated or clinically ineffective
- Do not use flavoxate, propantheline, trospium, fesoterodine, tolterodine MR or imipramine for the treatment of urinary incontinence (UI) or OAB
- Offer transdermal oxybutynin patches to patients unable to take oral medication
- There is no reason to expect patches or modified-release preparations of anti-cholinergic drugs to be more effective

4.1. Contraindications to antimuscarinic treatment

- Myasthenia gravis
- Significant bladder outflow obstruction
- Urinary retention
- Severe ulcerative colitis
- Toxic megacolon
- Gastrointestinal obstruction or intestinal atony

4.2. Caution with antimuscarinic treatment

Antimuscarinic treatment should be used with caution in the elderly (especially if frail), in those with autonomic neuropathy, hiatus hernia or reflux oesophagitis, and in those susceptible to angle-closure glaucoma.

Antimuscarinic treatment may worsen hyperthyroidism, coronary artery disease, congestive heart failure, hypertension, prostatic hyperplasia, arrhythmias and tachycardia.

Several commonly prescribed medications that are not within the anticholinergic class have significant anticholinergic effects, which when taken with known anticholinergic medication can increase the risk of adverse effects and have the potential to cause anticholinergic syndrome. These include:

- Antihistamines
- Tricyclic antidepressants
- Drugs for asthma and COPD
- Cold preparations
- Second generation antipsychotics (clozapine, olanzapine, quetiapine)
- Hyoscine

4.3. Antimuscarinic Syndrome

A confusional state with characteristic features related to dysfunction of the autonomic parasympathetic (cholinergic) nervous system. Symptoms classified into systemic and CNS manifestations:

- Systemic (peripheral) symptoms: blurred visions, photophobia, non-reactive mydriasis, loss of accommodation response, flushed and dry skin, dry mouth, tachycardia, hypertension and fever. Gastrointestinal and urinary motility are frequently reduced.
- CNS symptoms: delirium, agitation, disorientation and visual hallucinations. Ataxia, choreoathetosis, myoclonus and seizures may also occur without peripheral symptoms

5. Reviewing Overactive Bladder drug treatment

- Offer a face-to-face or telephone review 4 weeks after the start of each new OAB drug treatment. Ask the patient if they are satisfied with the therapy:
  - If improvement is optimal, continue treatment
  - If there is no or suboptimal improvement or intolerable adverse effects change the dose, or try an alternative OAB drug (see above), and review again 4 weeks later
  - Consider the use of objective measures such as bladder diaries, if feasible, to quantify the level of improvement
- Offer review before 4 weeks if the adverse events of OAB drug treatment are intolerable
- If second-line or third-line therapies are no more effective or tolerable than previous therapies revert to the previous and less expensive treatment and consider referral to secondary care
- Offer a further face-to-face or telephone review if a patient's condition stops responding optimally to treatment after an initial successful 4-week review
- Due to concerns around risk of cognitive impairment, falls and all-cause mortality associated with anticholinergic use, review patients who remain on

long-term drug treatment annually (or every 6 months for patients over 75 years).
  - Consider a 'drug holiday' for 4 weeks, and if successful discontinue treatment. Some patients will be able to manage their symptoms without long-term pharmacological therapy and have no further problems.
  - For those patients whose symptom control decline and were better managed whilst on treatment, restart.
- STOP anti-cholinergic drugs where the following is suspected or being investigated:
  - Dementia (increased confusion, agitation)
  - Chronic glaucoma (acute exacerbation of glaucoma)
  - Chronic constipation (exacerbation of constipation)
  - Chronic prostatism (urinary retention)
- If the patient wishes to discuss the options for further management (non-therapeutic interventions and invasive therapy) refer to a specialist secondary care centre to arrange urodynamic investigation to determine whether detrusor over-activity is present and responsible for OAB symptoms
  - If detrusor over-activity is present and responsible for the OAB symptoms, refer the patient back to the centre that conducted the urodynamic investigation if the patient would like to consider invasive therapy [note: referral to another centre will likely result in repeat tests and investigations being conducted as the information can be difficult to interpret]
  - If detrusor over-activity is not present refer to secondary care for further discussion concerning future management

6. Drug summary

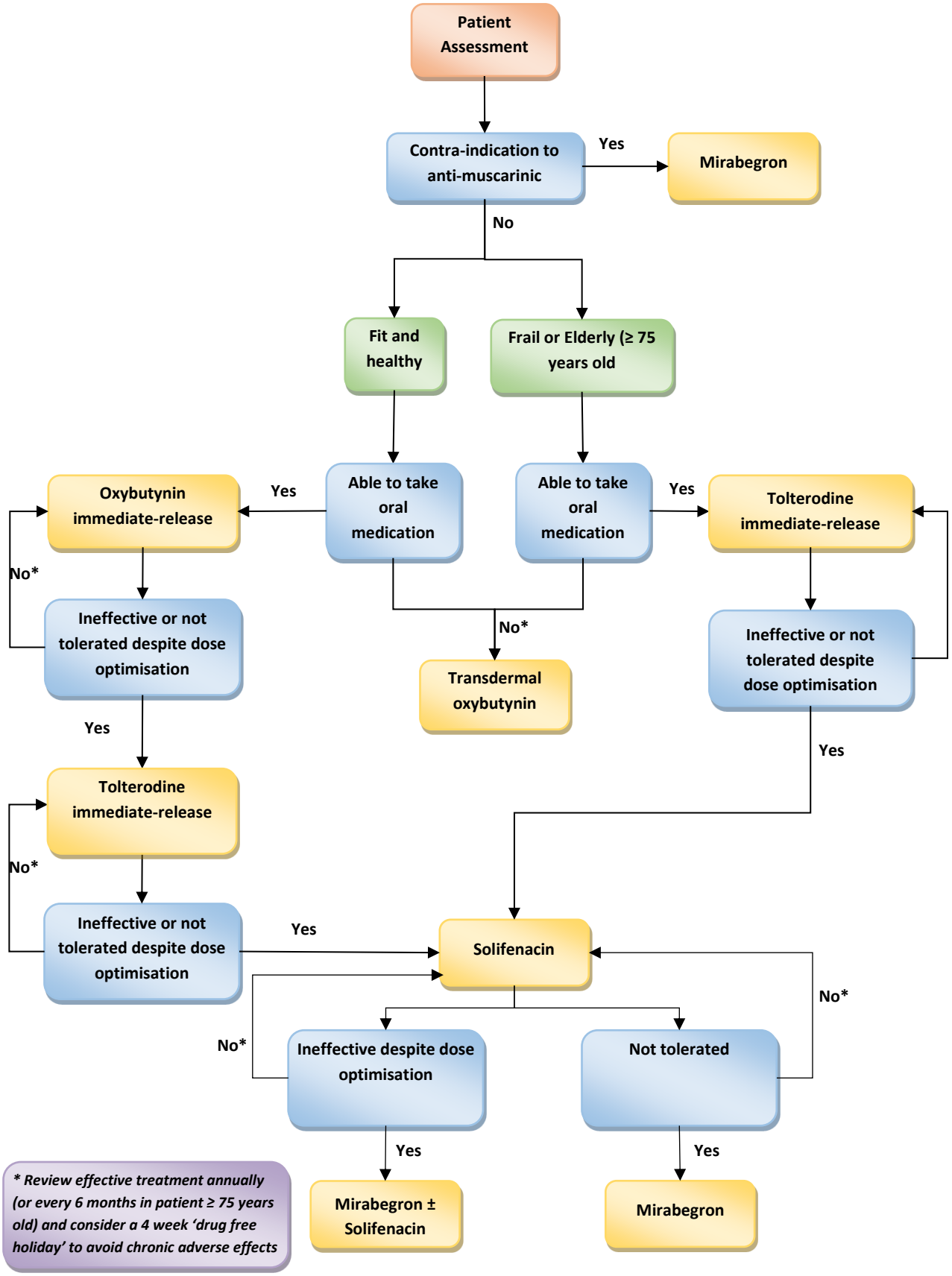
- If otherwise healthy: oxybutynin IR → tolterodine IR → solifenacin → mirabegron ± solifenacin
- If frail / elderly: tolterodine IR → solifenacin → mirabegron ± solifenacin
- If unable to swallow: transdermal oxybutynin
- If contra-indicated to an anti-cholinergic: mirabegron
- See Table 30 for quick reference to pharmacological therapies
- See Appendix 1 for summary flow diagram.

Table 30: OAB pharmacological therapy quick reference

Name	Initial dose	Maximum dose	Dosage adjustments*	Main side effects
Oxybutynin immediate-release	2.5mg - 5mg twice-daily or thrice-daily	5mg four-times daily	Nil	Dry mouth, constipation, blurred vision, dry eyes, cognitive impairment
Tolterodine immediate-release	2mg twice-daily	2mg twice-daily	Reduce to 1mg twice-daily if necessary to minimise side effects. Use with caution in	Dry mouth, constipation, blurred vision, dry eyes

Name	Initial dose	Maximum dose	Dosage adjustments*	Main side effects
			severe renal impairment.	
Solifenacin	5mg once-daily	10mg once-daily	Dose should not exceed 5mg daily in severe renal impairment or moderate hepatic impairment or those on potent CYP3A4 inhibitors. Should not be used in patients with severe hepatic impairment.	Dry mouth, constipation, blurred vision, dry eyes, low rate of cognitive impairment
Mirabegron	50mg once-daily	50mg once-daily	Dose should not exceed 25mg daily in moderate renal or hepatic impairment and those on potent CYP3A4 inhibitors. Should not be used in patients with severe renal or hepatic impairment.	Tachycardia, urinary-tract infection
Oxybutynin transdermal (Kentera)	1 patch applied twice-weekly to clean, dry unbroken skin on abdomen, hip or buttock (delivers 3.9mg / 24 hours)	1 patch applied twice-weekly	Rotate application site. Use with caution in patients with renal or hepatic impairment.	Skin irritation or pruritis, low incidence of dry mouth and constipation

Appendix 1: Summary of treatment recommendations



6.11 Qualitative Analysis (prescribing analysis)

6.11.1 Review of practice

Based on the large proportion of solifenacin prescribing identified from a review of HSCIC prescribing cost data, an analysis of change in practice was undertaken to determine the impact of the guideline and communication. Figure 49 identifies the change in solifenacin issued to patients in NHS Enfield from January 2014 following the above analysis (draft version). The reduction in proportion from 35% to less than 20% corresponds to a reduction in spend of £140,000 per annum. Unfortunately the other four CCGs within the NCL region have not followed through to the same extent however it identifies the opportunity that £700,000 of saving per annum exists within NCL alone. The data presented in Figure 50 and Figure 51 highlight the corresponding reduction in oxybutynin modified-release prescribing, and increase in oxybutynin / tolterodine immediate-release prescribing, respectively.

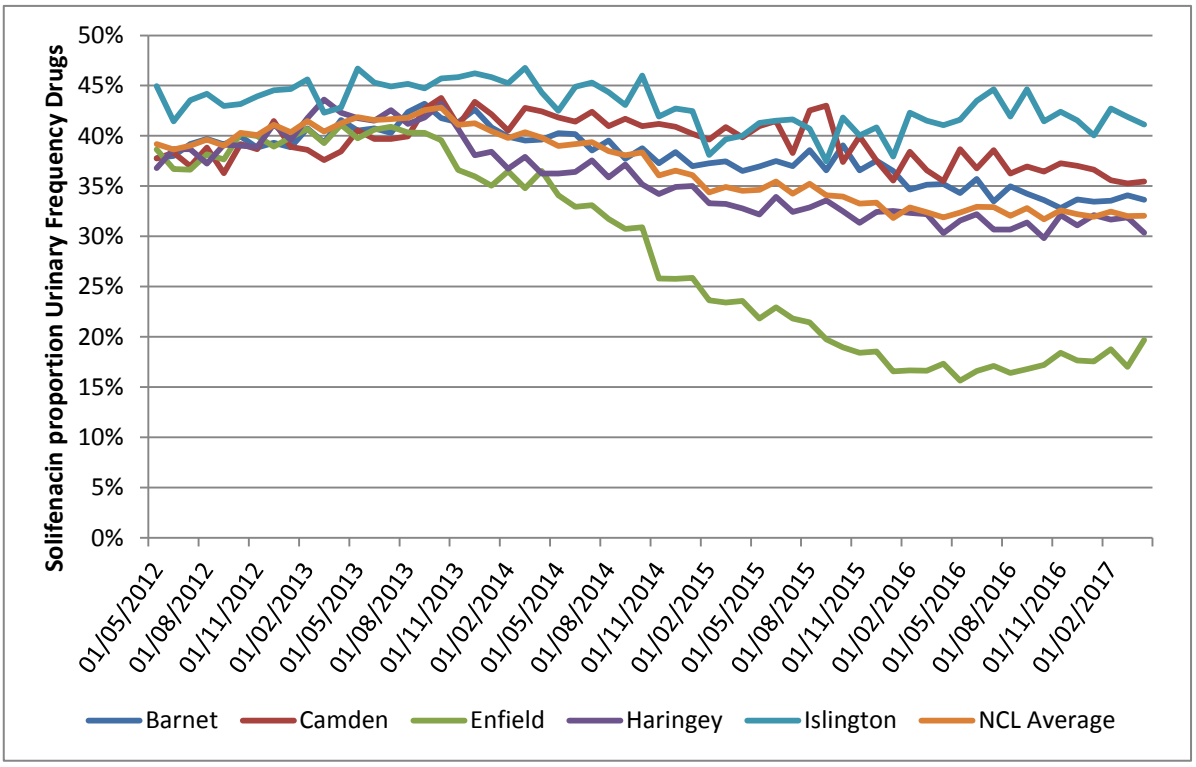


Figure 49: Proportion of solifenacin issued (as a proportion of all antimuscarinics) within North Central London between May 2012 and March 2017.

Practice begins to change following JfC discussion in January 2014; practice continues to modify further following draft guidelines in June 2014. Source: **https://openprescribing.net**

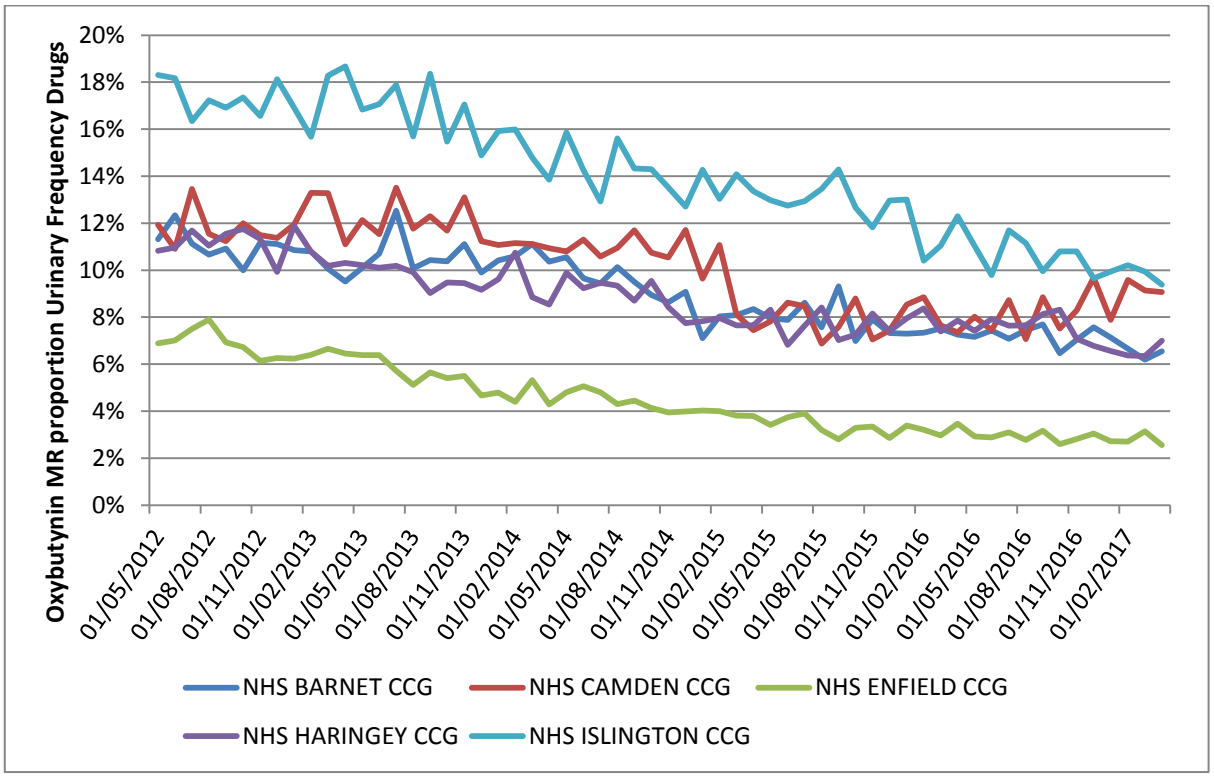


Figure 50: Proportion of oxybutynin modified-release issued (as a proportion of all antimuscarinics) within North Central London between May 2012 and March 2017.

Practice begins to change following JfC discussion in January 2014; practice continues to modify further following draft guidelines in June 2014. Source: **https://openprescribing.net**

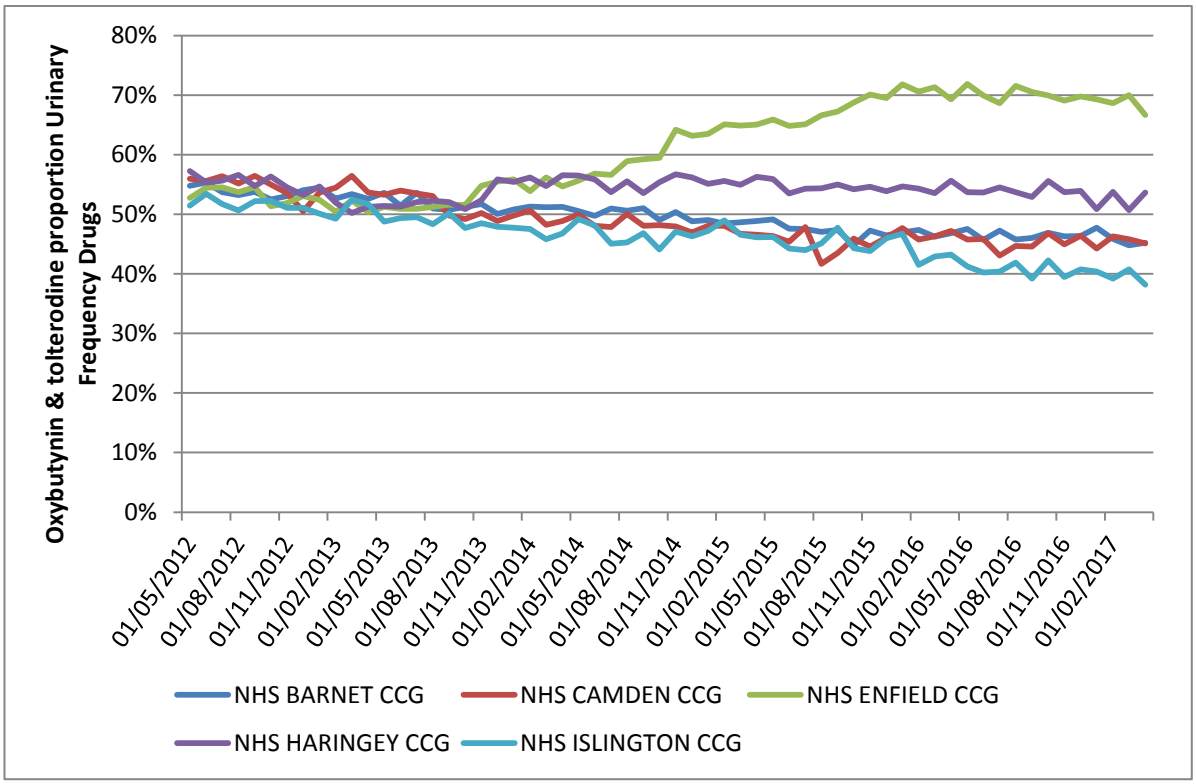


Figure 51: Proportion of oxybutynin and tolterodine immediate-release issued (as a proportion of all antimuscarinics) within North Central London between May 2012 and March 2017.

Practice begins to change following JfC discussion in January 2014; practice continues to modify further following draft guidelines in June 2014. Source: **https://openprescribing.net**

6.12 Reflection of this project

Further to the methodological work conducted previously (specifically the mixed treatment network meta-analysis for the antiepileptic drugs); it was considered important to ensure that this piece of work was not only a replication of methodological technique, but also to demonstrate growth in knowledge and analytical skills. To ensure this, two specific actions were taken as part of this research project:

1. Validation and improvement of NMA technique: The work was forwarded to Dr Nicola Welton (University of Bristol) for peer review. This choice was made as Bristol University is renowned for its School of Social and Community Medicine which is host to academic leads in Medical Statistics and Epidemiology. Dr Welton is an expert in the field of network meta-analysis and Bayesian multiparameter evidence synthesis, which this project was based on. Furthermore, Dr Welton is Deputy Director of the NICE Clinical Guidelines Technical Support Unit, which is based at the School, and is co-author on the two Technical Support Documents referenced within the methodology section of this project. Validation of the NMA model from an expert in the field and comments on further analyses which were originally missing (see section 6.9) provides assurance of an understanding of the methodology, interpretation of results, and overall robustness of project.

2. Translation to practice: The objective of this work was the development of a clinical guideline for the purpose of primary care clinicians, with the findings of the network meta-analysis shared with a Consultant Urologist who practices as part of a multi-disciplinary team (Urology, Uro-gynaecology and Care of the Elderly) where such treatment is frequently used. This was not done with the epilepsy group as part of research project one and therefore could be considered a limitation of the exercise. The particular choice of individual was made on the basis of familiarity with the processes of the Drugs and Therapeutics Committee, having conducted a number of submissions and being the recipient of challenging requests for information. The

inclusion of a respected clinician as part of the peer review, and known to the target audience, would ensure that the conclusion and recommendations carry value beyond an academic exercise.

[BLANK PAGE]

7.0 Chapter 7: Research Project Four – Oral anticoagulants for prevention of stroke in atrial fibrillation

7.1 Introduction

This fourth research project continues with the use of Bayesian NMA, with the trade-off analysis encompassing a number of endpoints, including efficacy, safety and cost. Through collaboration with the Medical Statistics and Epidemiology Division (School of Social and Community Medicine, University of Bristol), this project also includes a formal cost-utility analysis, bespoke to the UK setting, to advise UK NHS policy.

The use of a Bayesian meta-analysis rather than the traditional frequentist analysis is suitable for this project as again only a limited number of published trials have investigated these agents in a direct head-to-head comparison.

The goal of this project is to consolidate techniques developed over projects one to three and apply them to an analysis of a more complex therapeutic area that features as a priority for the local and national NHS service. The findings of the NMA and outcomes of the trade-off analysis would be used to influence prescribing across the UK. To ensure a robust and clinically relevant analysis based on learning from earlier projects, formal collaboration with a methodological, health economics and clinical expert team will be undertaken.

An additional goal for this project is to understand if any recommendations can be made on which of the agents under analysis should be prioritised to be trialled in a head-to-head manner.

7.2 Literature Review

7.2.1 Atrial fibrillation, stroke and myocardial infarction

Atrial fibrillation (AF) is the most common cardiac arrhythmia.(235) The prevalence of AF roughly doubles with each decade of age, rising to almost 9% at age 80-90 years.(236) AF substantially

increases (by up to five times) the risk of thromboembolic stroke (annual incidence 114 per 100,000) due to pooling of blood in the left atrium and systemic embolisation to the brain. More than 20% of the 130,000 annual strokes in England and Wales are attributed to AF. Approximately one-third of stroke patients die in the first ten days, one-third recover in one month and one-third have disabilities needing rehabilitation making stroke the leading cause of adult disability. Patients with thromboembolic stroke from AF have higher mortality, morbidity and hospital stay than patients with other stroke subtypes.

Since its launch in 1954 as a medicine for human use, warfarin remains an effective oral anticoagulant for the prevention of stroke in patients with AF,(237) however it is underused in clinical care given its cumbersome testing and monitoring requirements.

Although the incidence and mortality of stroke continue to fall in the UK, the underutilisation of anticoagulation in patients with AF at high-risk of stroke is a major gap in clinical care.(238) In patients with atrial fibrillation, antiplatelet and anticoagulant therapies are generally considered from the perspective of mitigation of stroke risk. However, the presence of atrial fibrillation is also associated with an approximately two-fold higher risk of future acute myocardial infarction,(239) whose annual incidence in England (130 and 55.9 per 100,000 for men and women respectively)(240) is higher than that for stroke.

7.2.2 Current usage and cost of warfarin in the NHS

A 2007 *Health Technology Assessment* report stated that approximately 950,000 people (2% of the GP population) in the UK are prescribed warfarin; increasing by about 10% per year.(241) Warfarin-related bleeding is one of the top five reasons for hospitalisation for adverse drug effects in England,(242) because of the narrow therapeutic index and numerous drug / dietary interactions. Although the approximate acquisition cost of warfarin is only £10 per patient per year, the requirement for therapeutic

monitoring to ensure optimal efficacy and to reduce the risk of bleeding, through hospital-, primary care-, or pharmacist-based anticoagulation clinics, or by home-monitoring with anticoagulant clinic support, increases the cost of warfarin treatment.

The estimated annual cost of managing patients on warfarin in the NHS in England and Wales is approximately £90 million. A 2006 NICE report estimated that 46% of patients who should be on warfarin are not receiving it, and that many receiving anticoagulation are not in optimal therapeutic range, perhaps because of concern about the risk and inconvenience of warfarin treatment.(243)

7.2.3 Novel oral anticoagulants and their place in the NHS

The class of novel (or non-vitamin-K antagonist) oral anticoagulants (NOACs) includes dabigatran (a direct inhibitor of clotting factor II), as well as rivaroxaban, apixaban, edoxaban, otamixaban and betrixaban (all factor X inhibitors). These agents have a more rapid onset and offset of action than warfarin, are therefore considered to have more predictable dosing requirements than warfarin, negating the need for therapeutic dose monitoring, increasing convenience and reducing overall cost.(244) However, the safety and efficacy of at least one of the NOACs (dabigatran) may vary according to achieved plasma concentrations, which may differ between individuals receiving the same dose,(245) suggesting a potential benefit from therapeutic dose monitoring (TDM). If this is proved to be the case, the consequence would be an increase in the overall cost of treatment.

Over recent years these drugs have been evaluated in clinical trials as an alternative to warfarin for the prevention of stroke in patients with AF (for which warfarin is given lifelong); as an alternative to LMWH for prevention of VTE in high-risk patients undergoing major orthopaedic surgery as well as those being hospitalised with acute medical conditions (for which LMWH is given to cover the high-risk period); as an alternative to a period of LMWH and then warfarin for acute treatment of a first VTE (usually for 6 months); as well as for secondary prevention after a first episode of VTE, for which there is currently no widely used treatment.

The estimated annual acquisition cost per patient of new anticoagulants is substantially higher than that of warfarin and will remain so until patent expiry (for example, 2020 for rivaroxaban). However, the higher acquisition cost could be offset by the reduced need for therapeutic monitoring through anticoagulation services, by increased effectiveness, or by improved safety. Potential limitations of NOACs include class- and drug-specific cautions / contraindications, potential for sub-therapeutic dosing, reduced adherence due to lack of regular monitoring, absence of, or limited experience with antidotes, as well as the added cost of maintaining stocks of numerous different anticoagulants and the potential for prescribing errors due to unfamiliarity.

7.3 Rationale for undertaking this research

Limitations of the previous evidence base (and shortfalls in previous attempts at evidence synthesis) make rational selection from the now wide range of available oral anticoagulants difficult for NHS commissioners, doctors and patients. Much of the existing NICE guidance in this area is limited to technology appraisals of the individual agents (for example, TA170 for rivaroxaban).

Clinical trials in this area have the following limitations:

- Few, if any, trials have made direct comparisons of NOAC drugs with one another. This is a recurrent theme in the licensing of new drugs as there is no regulatory requirement for the Pharmaceutical Industry to undertake this. This limitation may be addressed through the use of network meta-analysis to estimate the comparative efficacy and safety of agents, which have been tested against a common comparator; in this case, warfarin.
- Different rates of sub-therapeutic anticoagulation with warfarin within trials (as measured by the time spent in the therapeutic range) may have artificially inflated the apparent efficacy or deflated potential safety / tolerability concerns of the newer agents. This limitation may be addressed to some extent by investigating the relation of average time in therapeutic range (TTR) with efficacy within the network meta-analysis framework.

Prior synthesis research in this area has the following limitations:

- Some meta-analyses preceded the publications of potentially influential trials
- Failure to fully incorporate evidence on the adverse effects of oral anticoagulants by including data from all trials, regardless of indication, to maximise power and provide the most robust evidence on the balance between benefit and harm
- The lack of cost-effectiveness analyses relevant to England and Wales (i.e. the UK NHS setting)

Thus, it was agreed that there is a need for an up-to-date comprehensive evidence synthesis of all competing treatments to inform the rational choice of a minimum set of oral anticoagulants needed by the NHS for the main therapeutic indications to avoid unnecessary over-stocking and to reduce the risk of prescription error due to unfamiliarity. A working group between UCL and University of Bristol was created with a proposal drafted and submitted to the NIHR Health Technology Assessment (HTA) for funding. Following review by the HTA Clinical Evaluation and Trials Board, the research project was successfully awarded £242,550 (HTA – 11/92/17).

The work undertaken within this project also falls within the MRC ConDuCT (collaboration and innovation for difficult or complex randomised controlled trials) Hub, specifically *'evidence synthesis and expected value of information for prioritisation of RCTs'*.

7.3.1 Research aim

To determine what is / are the best oral anticoagulant(s) for prevention of stroke in atrial fibrillation; and for primary prevention, treatment and secondary prevention of venous thromboembolic disease. Note, the aim of this thesis relates only to stroke prevention in AF, as the area of the project that I was directly involved in.

7.3.2 Research objectives

The specific objectives of this research project were:

- To identify the most effective, safe and cost-effective anticoagulant for stroke prevention in AF, and consider whether the evidence is consistent across important patient subgroups (for example presence of comorbidities, age, etc.)
- To identify optimal anticoagulation strategies for use by Acute Trust Drugs & Therapeutics Committees or Area Prescribing Committees, based on the best drug(s) for each of the main therapeutic indications
- To estimate the value of conducting further research on the cost-effectiveness of these drugs, for example recommendation and identification of conducting a head-to-head trial of two or more new anticoagulants

7.4 Collaboration

Owing to the size of the review, a Project Board was established to undertake the NIHR funded Health Technology Assessment. The Project Board comprised of five key teams: (1) systematic review; (2) methodology; (3) clinical; (4) health economics; and (5) patient partner. Tasks undertaken by me relate to systematic review and clinical aspects, specifically: contribution to the pre-study protocol and grant applications; undertaking the systematic review, data extraction and assessment of data for the AF review; provision of clinical and pharmaceutical expertise to advise the health economics team; and supporting interpretation of the results in relation to clinical practice. The working group convened on a regular basis for 15 months over the course of the project. The execution of data derived from the systematic review within the Bayesian NMA was performed by the methodology team. The cost-effectiveness analyses were performed by the health economics team.

7.5 Protocol (Quantitative Analysis)

7.5.1 Systematic Review

A systematic review, with network meta-analysis, of randomised controlled trials addressing the question relevant to the study objectives was undertaken.

This review was undertaken in accordance with the Centre for Reviews and Dissemination (CRD) guidelines for undertaking systematic review,(246) and the Cochrane Handbook for Systematic Reviews of Interventions,(27) (as updated online during 2011: see www.cochrane-handbook.org). This review was prospectively registered in the PROSPERO database (http://www.crd.your.ac.uk/prospero), with registration number CRD42013005324. Related reviews were assigned registration numbers CRD42013005331 and CRD42013005330.

7.5.2 Search strategy

Scoping searches conducted during protocol development identified some previously published NMAs of oral anticoagulants. Studies included in these NMAs were rescreened against the eligibility criteria and a search was developed to identify any additional studies published beyond the search dates of the most recent NMAs in each population.(247)

The search strategy identifying studies for the review of stroke prevention in AF was run on 12 March 2014 and updated on 15 September 2014. The search strategy covered the period 2010 to September 2014 and included terms for AF, the interventions and comparators of interest and a filter added to focus the search on RCTs.

Relevant randomised trials were searched for within MEDLINE and PREMEDLINE In-Process & Other Non-Indexed Citations, EMBASE and The Cochrane Library. In keeping with the strategy of Chapter 6.0 (section 6.5.2), no restrictions on language were applied for this search. Duplicate records (identified by title, authors, journal citation and date published) were removed. The principal search strategy is outlined in Table 31.

Owing to the size of the team involved in this project, unlike the previous three areas, data from studies in progress, unpublished research and research reported in the grey literature from www.clinicaltrials.gov (to September 2012) was also obtained. The reference lists of retrieved studies and relevant review articles were screened. The NHS Economic Evaluation Database (NHS EED) and NICE Technology Appraisals records were searched.

7.5.2.1 Search terms

Table 31 provides a list of the search terms and limits applied when conducting the database search.

Table 31: Search strategy used in MEDLINE (in-process & other non-indexed citations – current week) and MEDLINE (1950 to present) for the analysis of NOACs for the prevention of stroke in AF (search date 20 March 2014; updated 15 September 2014)

	Search term	Results
01.	tachycardia, supraventricular/ or tachycardia, ectopic atrial/	5,440
02.	Atrial fibrillation/	33,510
03.	((atrial or atrium or auricular) adj3 fibrillat$).ti,ab.	38,980
04.	Heart fibrillat$.ti,ab.	42
05.	(supraventricul$ adj3 (arrhythmi$ or tachycardia$)).ti,ab.	7,547
06.	((atrial or atrium) adj3 (tachycardia$ or arrhythmi$)).ti,ab.	6,888
07.	(atrial adj3 tachyarrhythmi$).ti,ab.	1,210
08.	Atrial flutter/	4,944
09.	((atrial or auricular) adj3 flutter$).ti,ab.	5,382
10.	**Or/1-9**	**59,756**
11.	exp *anticoagulants/	94,278
12.	exp *coumarins/	24,265
13.	Warfarin/	14,307
14.	exp vitamin k/ai [antagonist and inhibitors]	1,534
15.	Thrombin/ai [antagonist and inhibitors]	3,372
16.	Factor Xa/ai [antagonist and inhibitors]	2,197
17.	Aspirin/	37,712
18.	(anticoagula$ or anti-coagula$).ti.	21,584
19.	(oral anticoagula$ or oral anti-coagula$).ti.ab.	7,768
20.	(coumarin$ or coumadin$ or warfarin or marevan or dicoumarol or dicoumarin or dicumarin or dicumarol or acenocoumarol or phenindione or aldocumar).ti,ab.	26,479
21.	(factor Xa adj2 (antagonist$ or inhibitor$)).ti,ab.	1,502
22.	(factor 10a adj2 (antagonist$ or inhibitor$)).ti,ab.	2
23.	(factor IIa adj2 (antagonist$ or inhibitor$)).ti,ab.	29
24.	((vitamin K or vitamin-k) adj2 (antagonist$ or inhibitor$)).ti,ab.	2,080
25.	(dabigatram or pradaxa or BIBR1048 or Apixaban or Eliquis or BMS-562247-01 or Edoxaban or Lixiana or savaysa or DU-176b or betrixaban or PRT-054021 or PRT0504021 or rivaroxaban or xarelto or BAY-59739 or Erixaban or D0913).ti,ab.	1,330
26.	(NOAC or NOACS).ti,ab.	152
27.	(aspirin or acetyl-salicylic acid or acetylsalicylic acid).ti,ab.	42,763
28.	**Or/11-27**	**175,520**

	Search term	Results
29.	**10 and 28**	**6,721**
30.	Letter/	829,317
31.	Editorial/	348,841
32.	News/	159,841
33.	exp historical article/	318,220
34.	Anecdotes as topic/	4,508
35.	Comment/	572,414
36.	Case report/	1,665,104
37.	(letter or comment$).ti.	94,907
38.	**Or/30-37**	**3,300,100**
39.	randomized controlled trial/ or Randomized Controlled Trials as Topic/ or random$.ti,ab.	835,720
40.	**38 not 39**	**3,270,043**
41.	Animals/ not humans/	3,807,926
42.	exp Animals, Laboratory/	714,413
43.	exp Animal Experimentation/	6,188
44.	exp Models, Animal/	407,073
45.	exp rodentia/	2,629,200
46.	(rat or rats or mouse or mice).ti.	1,097,935
47.	**Or/40-46**	**7,662,407**
48.	**29 not 47**	**5,201**
49.	Systematic review/	0
50.	Meta-analysis/	45,623
51.	(meta analy$ or metaanaly$ or metanaly$ or meta regression).ti,ab.	61,909
52.	((systematic$ or evidence$) adj2 (review$ or overview$)).ti,ab.	71,965
53.	(reference list$ or bibliograph$ or hand search$ or manual search$ or relevant journals).ab.	24,936
54.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.	26,492
55.	(search$ adj4 literature).ab.	26,789
56.	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or cinahl or science citation index or bids or cancerlit).ab.	82,698
57.	Cochrane.jw.	10,337
58.	((multiple treatment$ or indirect or mixed) adj2 comparison).ti,ab.	901
59.	**Or/49-58**	**186,127**
60.	randomized controlled trial.pt. or randomized controlled trial/ or Randomized Controlled Trials as Topic/	452,445
61.	controlled clinical trial.pt.	87,837
62.	Randomi#ed.ab.	343,274
63.	Placebo.ab.	151,447
64.	Drug therapy.fs.	1,674,296
65.	Randomly.ab.	208,182
66.	Trial.ab.	297,177
67.	Groups.ab	1,328,911
68.	**Or/60-67**	**3,312,451**
69.	clinical trials as topic.sh.	168,554
70.	Trial.ti.	123,158
71.	**Or/60-63,65,69-70**	**946,554**
72.	**59 or 71**	**1,075,719**
73.	**48 and 72**	**1,764**
74.	Limit 73 to yr=”2010-current”	728

7.5.2.2 Quality Assessment and Data Collection

Two reviewers independently screened the results of the searches by title and abstract (initially by me and verified by Mr Okoli). Disagreements were resolved through consensus or referral to a third reviewer where necessary (Ms Davies). The full text articles of all potentially relevant reports were obtained and assessed independently against the eligibility criteria. Multiple reports of the same study were collated and mapped under the primary publication.

Data extraction was undertaken against dedicated forms which were piloted initially on a small selection of studies to ensure their appropriateness. Data were extracted from the trial reports by me and checked by a second for all studies. Disagreements were resolved through consensus or by referral to a third reviewer where necessary (as above). The following data were extracted: study details (identifier, study design, location, year, length of follow-up, industry sponsorship); participant details (number of participants, age, gender); intervention details (drug name, dose, timing); comparator details; details relevant to risk-of-bias assessment (including adherence to and withdrawal from randomised allocation); and effect modifiers. Multiple reports from a study informed a single data extraction form. Data were extracted and managed using Microsoft Access® 2013 software (Microsoft Corporation, Redmond, WA, USA).

Data were extracted in dichotomous form based on the full randomised samples, as number of events in intervention and control groups and numbers of participants. Details of follow-up time was collected (e.g. participant-years in each treatment group) as well as estimates of hazard ratios (HRs) and their confidence intervals (CIs) when available.

7.5.2.3 Assessment of Risk of Bias

Risk of bias in studies was assessed using the Cochrane Risk of Bias Tool.(36) Assessments were carried out by me and independently checked by Mr Okoli. All disagreements were

resolved through consensus or by referral to a third reviewer where necessary, as per section 7.5.2.2.

7.5.3 Selection of data

7.5.3.1 Choice of Interventions

To perform the NMA each intervention group from each trial was allocated to a category. Each intervention category formed a 'node' within the treatment network. Separate nodes were created for different doses or frequencies of administration for the oral anticoagulants. The different forms of vitamin K antagonists were assigned to one node term (i.e. warfarin); however three sub-nodes under the term were created to account for INR range 2-3, INR range 3-4, and other. Due to the number of intervention categories (or treatment nodes), clear labelling was necessary.

7.5.3.2 Choice of Time Points

For trials where outcome data were presented at multiple time points the longest period of follow-up was taken.

7.5.4 Eligibility criteria

7.5.4.1 Study Design

Trials were included if they were of phase II or phase III randomised controlled design, and using either a superiority or non-inferiority design.

7.5.4.2 Participants

Adults (≥18 years) eligible for oral anticoagulation or antithrombotic treatment with a diagnosis of non-valvular atrial fibrillation (AF) were included. Trials in participants only eligible for parenteral (injected) anticoagulation were excluded. Unless otherwise specified, anticoagulation services may have been delivered in hospital, primary care, pharmacy-based clinics or through home monitoring and telephone support. The review was not limited to NHS anticoagulation services.

7.5.4.3 Inclusion criteria: Interventions and Comparators

Five oral NOACs were the focus of the review: dabigatran (Pradaxa®); apixaban (Eliquis®); edoxaban (Lixiana®); betrixaban (pending Marketing); and rivaroxaban (Xarelto®).

NOACs not considered within this review and reasons for exclusion were: erixaban (current stage of development was unclear); ximelagatran (withdrawn due to liver toxicity); darexaban (YM150) and AZD0837 (both discontinued in early development phase); LY517717 and letaxaban (TAK442) (no available information on clinical development for both); and otamixaban (parenteral formulation only in development therefore likely to be a treatment for VTE and not AF).

The aim of the systematic review was to inform the network meta-analysis (NMA), therefore the comparator interventions were determined to ensure that they would provide information on the relative effectiveness of the interventions of interest. Comparators were chosen based on the possibility of informing indirect evidence on the relative effectiveness of oral anticoagulants, and on the 'distance' of these comparators from our interventions of interest in the network, which relates to the likely increase in precision in the estimates of relative effectiveness of the oral anticoagulants.

Specific comparators for the prevention of stroke in AF were:

- therapeutic doses of warfarin or other VKA [with optimal international normalised ratio (INR) range 2–4]
- aspirin
- clopidogrel

7.5.4.4 Exclusion criteria

Studies evaluating fixed-dose administration of warfarin were excluded. Studies evaluating warfarin with suboptimal target INR compared with UK guidelines were also excluded from the main analyses but combined with studies evaluating warfarin with standard target INR in sensitivity analyses.

7.5.5 Outcome measures

The key outcomes addressed in the NMA are marked with an asterisk (*) in the list below. These were chosen based on three considerations: (1) their clinical importance; (2) the consistency or reporting across studies included in the network; and (3) the number of data that were available for inclusion in NMAs. The choice was based on agreement by the clinical contributors to the research project (Prof Hingorani, Dr Sofat, and I).

Where outcome data were not presented directly, they were computed or substituted, using data for other outcomes using assumptions that were considered to be clinically reasonable, as guided by the clinical contributors to the research project. For example, where data could not be extracted for outcome ‘stroke or SE’, the data for ‘all stroke’ was used. Where CRB was not reported but both major bleeding and CRNM bleeding events were, the total number of events across these categories was used. This was performed by me and the data verified by Mr Okoli.

7.5.5.1 Outcomes: Prevention of stroke in Atrial Fibrillation

Data were sought on the following outcomes:

- Stroke or systemic embolism (SE)*
- All stroke
- Ischaemic stroke (major ischaemic stroke or minor ischaemic stroke)*
- Fatal stroke
- Non-fatal stroke
- Haemorrhagic stroke (major haemorrhagic stroke or minor haemorrhagic stroke)
- Any bleeding
- Minor bleeding
- Major bleeding*
- Clinically relevant non-major (CRNM) bleeding
- Clinically relevant bleeding (CRB)* (defined as CRNM bleeding or major bleeding)
- Intracranial bleeding*

- Extracranial major bleeding
- Extracranial minor bleeding
- Fatal bleeding
- Bleeding from surgical site
- Thrombocytopenia
- Myocardial infarction (MI)*
- Transient ischaemic attack (TIA)
- Arterial event
- Quality-of-life outcomes
- Hospital admission
- Death (cardiovascular)
- All-cause mortality*

7.5.6 Quantitative (Statistical) Analysis

For each analysed outcome in each review, a frequentist meta-analysis of 'direct evidence' and a Bayesian NMA of 'direct and indirect evidence' were performed. Results from individuals studies are presented in forest plots, arranged within each possible pairwise analysis (e.g. one NOAC versus another NOAC).

7.5.6.1 Undertaking the Quantitative Analysis

The networks were plotted to illustrate the data structure for each review and outcome. In these plots, the size of the node for each intervention is proportional to the number of patients randomised to that intervention. When direct evidence comparing two interventions was available, these two interventions are connected by a line where its thickness is proportional to the number of patients who contributed to the comparison. The intervention labels were formatted as follows:

- Licensed doses of NOACs are written in bold typeface; these are interventions of primary interest.
- Interventions that were excluded from the primary analysis labels are presented in square brackets. Such exclusions are because (1) they were not considered to be of interest to inform health decisions in the UK (e.g. warfarin interventions using sub-

therapeutic INR ranges); or (2) the total number of events was zero, so they are uninformative; or (3) they do not connect with the other trials in the network.

- Excluded interventions that were included in sensitivity analyses are marked with an asterisk (*) under section 7.5.5.1.

In previous chapters of this thesis the *random-effects* model of the Frequentist meta-analysis was used, however for this analysis the *fixed-effect* model was used. This was agreed by the Project Board following a review of preliminary analyses focussing on heterogeneity concluding that trial protocols and patient groups were sufficiently similar across all studies owing to the nature of strict diagnosis driving the inclusion criteria.(49)

All meta-analyses were performed within a Bayesian framework, using freely available WinBUGS software version 1.4.3 (MRC Biostatistics Unit, Cambridge, UK) and code.(248) As in section 6.0, the MTC model was coded for binomial likelihood owing to the endpoints of interest being discrete and not continuous. In supplementary analyses for some outcomes it was agreed that hazard ratios (HRs) would be modelled rather than odds ratios (ORs) as this more accurately describes the outcomes of risk being non-cumulative.

Convergence of the Markov chains was assessed as in previous chapters via visual examination of trace plots. Goodness of fit was also assessed by calculating the posterior mean residual deviance.

7.5.6.2 Investigation of heterogeneity

Subgroup analyses were planned to examine the extent to which patient- and study-level characteristics explain between-study heterogeneity. Important characteristics were pre-specified as age, gender, ethnicity / race, body mass index (BMI) or weight, renal status or creatinine clearance, blood pressure, diabetes mellitus, hypertension, previous thrombotic event, liver disease, chronic heart failure, cancer, pregnancy, intervention dose, average TTR in the

warfarin group, and summary assessment of risk of bias for each outcome.

Clinically relevant factors for AF trials were CHADS2, CHADS2 VASC, HAS-BLED, history of previous stroke or TIA and previous MI.

It is expected that investigation of between-study variation using these characteristics could not be studied in most cases due to the lack of multiple trials of the same pairwise comparison, although some sensitivity analyses for the review of stroke prevention in patients with AF could be conducted.

7.5.6.3 Investigation of inconsistency

The validity of a NMA depends on the assumption that there is no effect modification of the pairwise intervention effects or, that the prevalence of effect modifiers is similar in the different studies. For a clinical and epidemiological judgement of the plausibility of this assumption an assessment was undertaken to determine whether or not the trials were similar in ways that might modify treatment effect, based on the pre-specified list of potential effect modifiers (see section 7.5.6.2).

Direct and indirect estimates generated from the analysis will be presented for comparison. The extent of the disagreement between the direct and indirect estimates can be used as a local measure of inconsistency for that comparison. [Note that for the vast majority of comparisons there was either only direct evidence or only indirect evidence, so that the NMA estimates correspond to one of these]

7.6 Data Extraction

A bespoke database was created using Microsoft Access (Microsoft® Office Access 2003, Microsoft Corporation) to ensure that data extraction was undertaken in a systematic manner. The process of data extraction is described in section 10.4 (Appendix IV) and section 10.5 (Appendix V).

7.7 Protocol (Cost-Effectiveness Analysis)

The following section was led by the health economics working group with input from the clinical team in identifying key assumptions and conditions.

7.7.1 Research Questions

The clinical and health economics working groups of the project board had particular interest in identifying the answer to the following question: ***"What is the most cost-effective first-line anticoagulant in the prevention of ischaemic stroke in patients with AF?"***

The aim was to evaluate cost-effectiveness from an NHS perspective. Costs and outcomes were to be modelled over the expected lifetime of patients. The models were designed to synthesise evidence on a number of parameters [e.g. incidence of ischaemic stroke, relative treatment efficacy, adverse events (AEs), costs, etc.] in order to estimate the relative cost-effectiveness of treatment options. The 'model inputs' are based on a variety of evidence sources. These include routine data on drug costs and observational studies of the long-term costs and quality of life (i.e. utilities) in AF.

The AF model was constructed in R version 3.02 (The R Foundation for Statistical Computing, Vienna, Austria)(249) and performed by the health economics team with input from the clinical team. All NMA were conducted in WinBUGs.(58)

7.7.2 Patient Population

The clinical team considered patients with non-valvular AF who were eligible for anticoagulation. No distinction was made between paroxysmal, persistent and permanent AF. The RCTs identified in the systematic review did not distinguish between AF type; however, patients with paroxysmal AF are less likely to be included in RCTs than those with other AF types. The results from the cost-effectiveness analysis are therefore most applicable to patients with persistent and permanent AF.

The patients considered by the model were those receiving first-line anticoagulation at the age of 70 years, based on the mean age

observed in the RCTs identified in the systematic review [mean age 70 years, standard deviation (SD) 8 years], and consider costs and benefits over a lifetime. As assumption of a 60:40 split in favour of males was made, as indicated by recruitment into the RCTs.

7.7.3 Interventions

The first-line treatments for AF included within the cost-effectiveness analysis (CEA), alongside their standard or licensed doses, are listed in Table 32. To ensure relevance to the NHS, only licensed treatments and doses within the UK were included within the analysis. Although a few small RCTs have compared betrixaban with warfarin in AF, there was not enough evidence to include it in the economic model. Standard of care for patients with AF, before the introduction of NOACs, was warfarin, and remains the comparator.(250)

The clinical team (based on current practice) agreed that treatment switching may occur as a result of treatment failure, indicated by ischaemic stroke or serious AEs, such as intracranial haemorrhage (ICH). It was also agreed that for patients on warfarin as first-line treatment, the only second-line intervention available was assumed to be no treatment. In practice this would mean that for patients on a NOAC first-line treatment, second-line treatment may be either warfarin or no treatment, with no treatment being the only third-line option. These rules are illustrated in Figure 52, where the events that may lead to treatment switching are indicated.

7.7.4 Outcomes

Results on total costs and quality-adjusted life-years (QALYs) are presented, both discounted at 3.5%, as agreed by the clinical and health economics teams. This discount is made on the basis that future costs and benefits will be valued less by society in the future than the value at present day. The use of a uniform rather than differential (or variable) discount over a future time period as well as the value of 3.5% is commonly used for economic modelling in developed countries such as the UK and USA. These modelling terms are currently used by NICE as part of their Technology Appraisal program, based on the recommendations of the UK Treasury.(251)

The health economics team undertook probabilistic analyses, for which model parameters are given probability distributions to reflect uncertainty in their values.(252) The results are summarised with the expected costs, expected QALYs and expected net monetary benefit (NMB) for a range of willingness to pay per additional QALY gained (where expected values are an average over the joint distribution of the model parameters). The outcomes are aligned with NICE's willingness-to-pay threshold of £20,000–30,000 per QALY.(72)

Table 32: First-line NOAC and licensed doses (used to inform the CEA); Source – BNF 67 (229)

Intervention	Dose / Target INR	Time on treatment
Apixaban	2.5mg twice daily (elderly) 5mg twice daily	Lifetime
Dabigatran	110mg twice daily (elderly) 150mg twice daily	Lifetime
Rivaroxaban	20mg once daily	Lifetime
Warfarin	Variable dose to INR 2-3	Lifetime

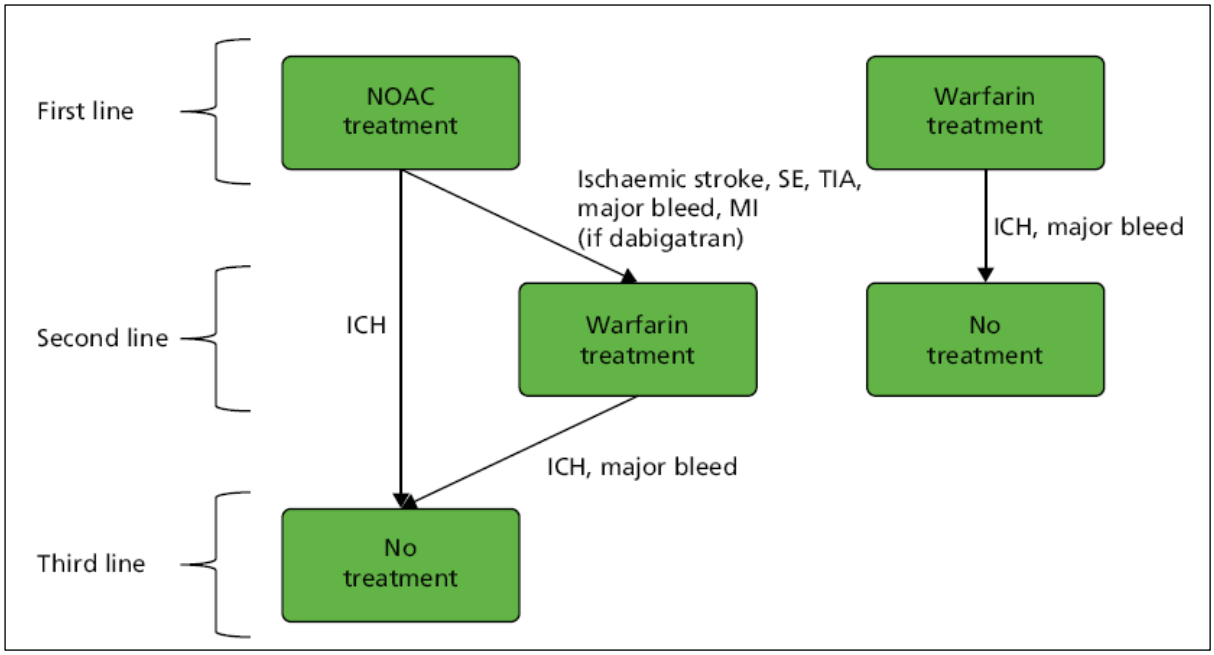


Figure 52: Treatment strategies and switching / discontinuation rules (events that may lead to treatment switching are indicated next to the arrows)

7.7.4.1 AF Model Structure

[This section of the research project was performed by the health economics team and is therefore included for information only]

The discrete-time Markov multistate model structure (see Figure 53) used a cycle length of 3 months, as in other recently published models.(253-255) The model was run for a cohort starting at age 70 years using a lifetime time horizon with a cut-off at 100 years. Patients were initially assigned to first-line treatment, which may be warfarin or a NOAC. There is a probability of switching to another therapy or discontinuing treatment entirely (see Figure 52).

Each of the treatment strategies has the same model structure but with different costs, utilities and event probabilities. From any state, a patient can have a clinically relevant (extracranial) bleed, an ICH, an ischaemic stroke, a MI, a TIA, a SE, can discontinue or switch treatment because of these events or die. These events are similar to those used in earlier published models.(255, 256) The primary difference to previous models is that this analysis does not distinguish between minor and major ischaemic stroke, as there was limited evidence from the RCTs to estimate the relative rates of these events. Non-clinically relevant minor bleed events were also not included as it is assumed that this will not have a significant impact on costs, quality of life or future risks. As in most previous models, memory states are used to record a history of the most important previous events.

The model assumes that SE and TIA have only short-term effects on future risks, costs and utilities, whereas ischaemic stroke, ICH, other CRB and MI have long-term consequences that must be modelled. Up to four major events are therefore recorded and assumed to affect future risks, costs and utilities. For example, patients with MI + ICH will have different risks, costs and utilities to patients with MI or ICH alone. Unlike the Wisloff 2013 model,(257) the current model does not distinguish between bleed locations, such as gastrointestinal and other types of bleed. Based on agreement from the clinical team members, it was assumed that the greatest impact on risks, costs and effects is captured by the broad

definition of ‘clinically relevant bleeds’, as reported in the RCTs. In total, the current model has 17 states, including a well state (‘AF Well’) and death.

At any cycle, patients can switch treatments to second-line or no treatment. All adverse health events increase the probability of treatment switching. An ICH is assumed to always lead to treatment switching. Patients are assumed to always switch treatment from dabigatran to warfarin if they experience a MI as a result of recent findings suggesting a link between dabigatran and MI risk.(258) Whether or not patients switch treatment after an ischaemic stroke depends on whether it was due to treatment failure or non-compliance. It was assumed that this was due to treatment failure but that only a proportion of patients will switch treatment following an ischaemic stroke.

In the Markov model, future state transitions depend only on the current state in which the patient is (and not past history). However, the age of the cohort will increase with each cycle and mortality risk increases accordingly, based on general population life tables. There is no available evidence to suggest treatment effects change with age or that they depend on event history. The model therefore makes the assumption that treatment effects are independent of age and event history.

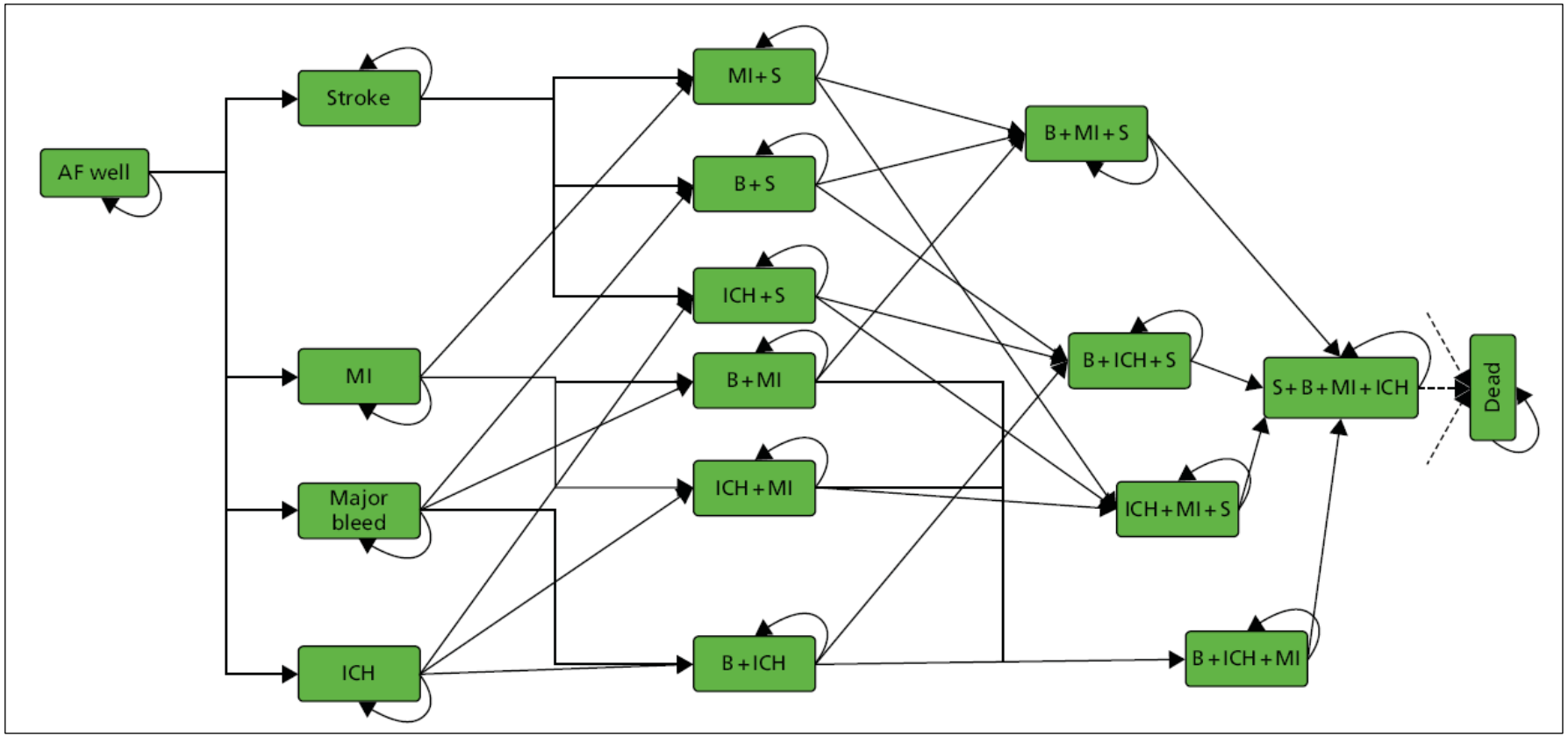


Figure 53: Markov model for AF. Patients can experience transient events (TIA or SE) but stay in the same health state, with possibly changed treatment, thereafter; (B) other clinically relevant bleed; (S) ischaemic stroke

7.7.5 Inputs for the CEA model for AF

7.7.5.1 Cost of Pharmacotherapy

[Guidance on the inputs for the following section of the research project was provided by me to the health economics team]

To ensure relevance to the NHS the costs of medicines were based on the current version of the British National Formulary (BNF) at the time, March 2014, edition 67.(229) The most economical pack sizes were used to inform on the price. Although in practice various areas of the NHS receive medicines at discounted prices against the BNF reference cost (particular in secondary care), these discounts are confidential and are not usually applicable in the primary care setting where the majority of prescribing and supply of these medicines for AF is delivered. As these are all under patent the validity of the model will hold true for a number of years.

Edoxaban does not currently have a list price within the UK, therefore for the base case analysis it was assumed that it would be price matched to dabigatran. As all of the NOACs are taken orally, it was assumed that there are no administration or monitoring costs. This follows the costing report in AF of Ali et al.(259) Average drug and monitoring cost of warfarin comes from a costing report by NICE68 and is cited in the study by Kansal et al.(255)

The unit costs of drugs are assumed to be fixed as they are known, so that point estimates, rather than distributions, were entered into the models. Given that the administration and monitoring costs of warfarin are variable (dependant on setting and level of monitoring required) and therefore present a level of uncertainty, in the absence of other information a uniform distribution was assumed ranging from 50% to 150% of the costs specified within the NICE costing report.(243) A sensitivity analysis for the assumed cost of warfarin monitoring was also performed.

Table 33: Drug dose, duration and costs for the prevention of stroke in AF

Intervention	Dose per day (mg)	Mg per tablet	No. in pack	Cost (£) per pack	Cost (£) per day	Administration cost (£)	Cost (£) per 3-month cycle AF model
Apixaban	10 5	5 2.5	56 60	61.50 65.90	2.20 2.20	0.00 0.00	200.42 200.44
Dabigatran	300 220	150 110	60 60	65.90 65.90	2.20 2.20	0.00 0.00	200.44 200.44
Rivaroxaban	20	20	60	210.00	2.10	0.00	191.63
Warfarin	Variable	Variable	100	Variable	Variable	sensitive	105.13*

* Values were inflated from the 2013-14 Office for National Statistics (ONS) Consumer Price Inflation Index for Medical Services (DKC3) and placed in a uniform distribution ~(52.57, 157.70) and (210.26, 630.79) on the cost per 3-month and annual cycles, respectively.(260)

7.7.5.2 Cost of Atrial Fibrillation and Anticoagulant-related events (Acute Phase)

All costs of acute and chronic care were inflated to 2013–14 values using the Office for National Statistics (ONS) Consumer Price Inflation Index for Medical Services (DKC3).(261) Acute management costs for SE, MI, TIA, and CRB come from the 2012–13 NHS reference costs.(262) The reference costs for MI account for only direct hospitalisation; we assumed that total costs would be double this amount to account for follow-up costs.(263) Acute management costs for ischaemic stroke and ICH come from a study of patients with AF on a UK stroke registry.(264) The health economics team assumed a normal distribution for the mean acute costs, with SDs defined by the standard errors of the source data.

7.7.5.3 Cost of Atrial Fibrillation and Anticoagulant-related events (Chronic Phase)

Long-term management costs of stroke (ischaemic stroke or ICH) also come from the UK stroke registry (264). This registry stratified the severity of ischaemic strokes by disability (non-disabling, moderately disabling, totally disabling); their annual costs and SDs were averaged, weighted by the number of events. As in the study by Kansal et al.,(255) the same cost for ICH as for ischaemic stroke were assumed. Normal distributions are assumed, with SDs defined by the standard errors of the source data.

7.7.5.4 Utilities

The AF model used utility weights combined with survival to estimate QALYs. Utility weights are anchored at 1 (best health) and 0 (as bad as death), such that a year spent in an intermediate health state with a utility weight of 0.5 would be considered equivalent to 6 months in the best health state with a utility value of 1. The models have a number of acute health events that affect patients for a short period, followed by a partial or full recovery and a number of chronic health states from which patients do not recover.

Utilities were identified from a previous NICE technology appraisal submission on rivaroxaban (253). The rivaroxaban technology appraisal submission conducted a systematic literature search for evidence on EQ-5D (European Quality of Life-5 Dimensions) utility index in health states related to AF.

7.8 Findings (Quantitative Analysis: Stroke prevention in AF)

7.8.1 Systematic Review

A total of 1,852 records were identified from the various databases for the systematic review of stroke prevention in AF (see Figure 54).

Twenty-three completed eligible RCTs were identified for inclusion in the review, with a total of 41 associated references for these trials. No ongoing trials were identified. A summary of the characteristics of the 23 trials is presented in Table 34.

Twenty of the trials were of multicentre design; two trials were conducted in two centres; and one trial was conduction in only one centre. The majority of the multicentre trials were conducted across several countries in North and South America, Europe, Asia, Russia and Israel, Australia and South Africa. The two-centre trials were conducted in one country: one in China and the other in Denmark. The single-centre trial was conducted in Denmark. Sixteen of the trials were Phase III studies and seven were Phase II studies. The number of patients randomised across the 23 trials ranged from 75 to 21,105, with a total of 94,656 patients, of whom 97% (91,333) were from the Phase III studies. Thirteen studies: six Phase III and seven Phase II studies examined a NOAC. Four studies examined edoxaban, three each examined apixaban and dabigatran, two examined rivaroxaban and one examined betrixaban.

Eligibility criteria for patient participation were similar across studies: all patients having non-valvular AF, whether new or existing, and including paroxysmal, persistent or permanent types. Diagnosis of AF was predominantly by electrocardiography. In a few cases, Holter recording, pacemaker or other intracardiac recording was used. This was important in reducing heterogeneity and the basis of using a fixed-effect model for the meta-analysis.

The mean age of included patients was reported in only 61% of the studies and this ranged from 63.3 to 81.5 years. The percentage of male patients was reported in 78% of the studies, and this varied significantly across the studies, ranging from 44.9% to 82.9%. Mean BMI was not often reported and ranged from 24.4 to 30.5 kg/m^2. Percentage of patients with previous stroke, hypertension and chronic heart failure varied significantly across the studies, ranging from 5% to 63.8%, from 38% to 93.7%, and from 0% to 100%, respectively. Bleeding risk among patients was assessed predominantly with the CHADS2 scoring system.

Warfarin was examined in all but two of the 23 of the included studies, against a NOAC in 12 studies, and against aspirin in nine studies. Standard intensity warfarin (INR 2–3) was examined by all of the studies, although in a few studies the warfarin arm was a mixture of low intensity (INR < 2) and standard intensity, in unknown proportions. Across all of the studies, mean TTR for warfarin ranged from 45.1% to 83% of the treatment duration. One study(265) compared both low intensity warfarin (INR < 2) and standard intensity (INR 2.5–3.5) dicoumarol with aspirin, but the mean TTR was not reported for the standard intensity dicoumarol arm. The doses of NOACs examined were edoxaban 30 mg, 45 mg, and 60 mg once daily and 30 mg and 60 mg twice daily; apixaban 2.5 mg and 5 mg twice daily; dabigatran 50 mg, 110 mg, 150 mg and 300 mg twice daily; rivaroxaban 15 mg and 20 mg once daily; and betrixaban 40 mg, 60 mg and 80 mg once daily. Examined aspirin dosages ranged from 75 mg to 325 mg once daily.

Treatment duration in the edoxaban and dabigatran studies was predominantly 3 months, although one study reported mean treatment durations of 24 months and another reported a median treatment duration of 29.8 months. Mean treatment duration for apixaban studies ranged from 13.1 to 21.6 months and one study reported 3 months' treatment duration. The two studies on rivaroxaban reported 30 months' treatment duration and a mean treatment duration of 19.4 months, respectively. Mean treatment duration 4.9 months was reported in the betrixaban study. Treatment duration was similar for each comparator in almost all the NOAC

studies. Reported efficacy and safety outcome types were similar across studies and these were reported at the end of the treatment periods. All 23 studies reported data on stroke, 15 studies reported data on MI, 18 studies reported data on major bleeding, 12 studies reported data on CRB, and 18 studies reported data on all-cause mortality.

Fifteen of the 23 studies, including all the NOAC studies, were sponsored by pharmaceutical companies. Six studies were funded by grants from medical research bodies although two of these grants contained contributions from a pharmaceutical company. Sponsor detail was not reported in two studies. In most of the pharmaceutical company sponsored studies, the sponsor(s) had influence on the study design, data management and analysis.

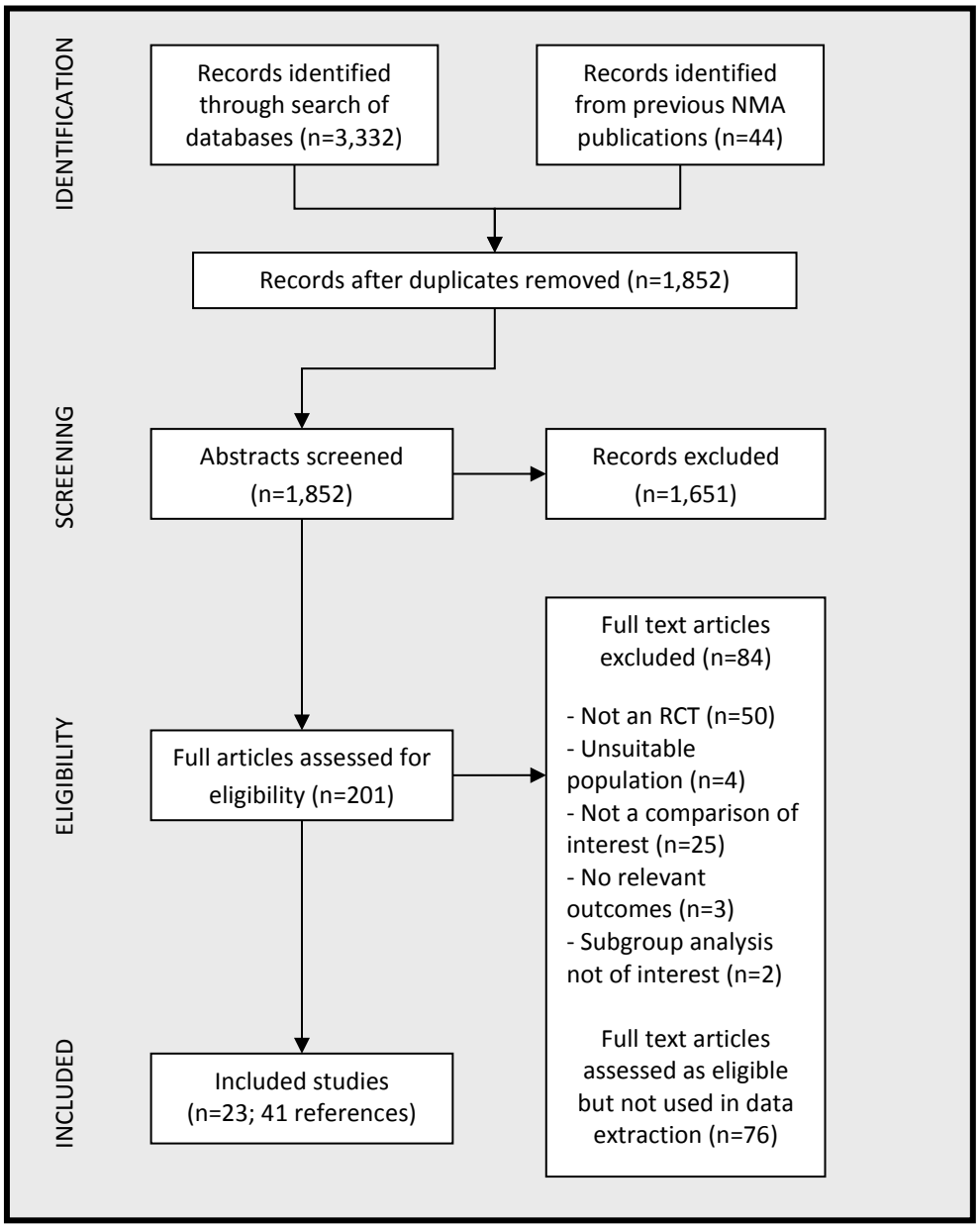


Figure 54: The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of studies identified from the systematic review for inclusion in the meta-analysis for stroke prevention in AF

Table 34: Characteristics of randomised controlled trials of NOACs included in the quantitative analysis of stroke prevention in AF

Study (centre type) [countries]	Study type, Sponsor (Sponsor involvement)	Age, eligibility (mean age, years) [% male]	AF type	No. randomised	Interventions compared	Treatment duration (months)	Mean TTR (INR), %	Outcomes	Time of outcome assessment (months)
ACTIVE W (265) (Multi-centre) [North and South America, Europe, Russia, Israel, Australia, Asia and South Africa]	Phase III Sanofi Aventis and Bristol Myers Squibb (The sponsor contributed to the study design 'but had no role in data collection, data analysis, data interpretation, or writing of the report')	≥ 18 (70.2) [66.1]	Non-valvular ECG diagnosed	6,706	Antiplatelet: 1. Clopidogrel 75mg + aspirin 75-100mg once daily Warfarin: 2. INR 2-3 (an equivalent VKA may have been used)	Not stated	63.8	**Efficacy:** all stroke, ischaemic stroke, haemorrhagic stroke, MI **Safety:** all bleeding, major bleeding, minor bleeding, fatal bleeding, death (all causes)	15.4
AFASAK (266) (Two centres) [Denmark]	Phase III NycoMed AS; Henrik Henriksen's Foundation; Kathrine and Vigo Skovgaard's Foundation; and Danish Medical Research Foundation (not stated)	≥ 18 (74.2) [53.6]	Chronic non-valvular ECG diagnosed	1,007	Warfarin: 1. INR 2-3 Antiplatelet: 2. Aspirin 75mg daily 3. Placebo once daily	24	73	**Efficacy:** all stroke, fatal stroke, minor ischaemic stroke, TIA **Safety:** bleeding, death (all causes)	24

AFASAK II (267) (Single centre) [Denmark]	Phase III Danish Heart Foundation; NycoMed DAK; DuPont Pharma; Danish Foundation for Medical Research; other non-industry funders (not stated)	≥ 18 (74.2) [60]	Chronic non-valvular ECG diagnosed	677	Warfarin: 1. 1.25mg/day fixed dose 2. 1.25mg/day fixed dose plus aspirin 300mg/day 3. INR 2-3 Antiplatelet: 4. Aspirin 300mg daily	42	73	**Efficacy:** all stroke, ischaemic stroke, haemorrhagic stroke, fatal stroke, stroke or SE, TIA, MI **Safety:** major bleeding, minor bleeding, intracranial bleeding, death (all causes)	42
AF-ASA-VKA-CHINA (268) (Two centres) [China]	Phase III Grant from talent pool subject of Shanghai Shi Dong Hospital (not applicable)	≥ 80 (NR) [NR]	Persistent and permanent non-valvular Confirmed by the case history and ECG	110	Warfarin: 1. INR 1.6-2.5 Antiplatelet: 2. Aspirin 100mg daily	24	NR	**Efficacy:** stroke or SE, ischaemic stroke, MI **Safety:** all bleeding, major bleeding, minor bleeding, fatal bleeding, death (all causes)	1, 6, 12, 18, 24
AF-DABIG-VKA-JAPAN (269) (Multi-centre) [Japan]	Phase II Boehringer Ingelheim (The sponsor was involved in the trial)	≥ 20 (NR) [NR]	Paroxysmal persistent or permanent non-valvular ECG diagnosed	174	Dabigatran: 1. 110mg bd 2. 150mg bd Warfarin: 3. INR 2-3 (INR ≥ 1.6 to ≤ 2.6 in ≥ 70 years)	3	NR	**Efficacy:** stroke or SE **Safety:** all bleeding, major bleeding, composite CRB	3

AF-EDOX-VKA-ASIA (270) (Multi-centre) [Taiwan, South Korea, Hong Kong and Singapore]	Phase II Daiichi Sanyo Co. (The sponsor 'had influence on the study design, data management and analysis, and key decisions')	18-80 (65.1) [65.4]	Non-valvular ECG diagnosed $CHADS_2 \geq 1$	235	Edoxaban: 1. 30mg daily 2. 60mg daily Warfarin: 3. INR 2-3	3 (edoxaban) 6 (warfarin)	45.1	**Efficacy:** stroke or SE **Safety:** all bleeding, major bleeding, minor bleeding, CRNM bleeding	3
AF-EDOX-VKA-JAPAN (271) (Multi-centre) [Japan]	Phase II Daiichi Sanyo Co. (The sponsor 'had input on the study design and data analysis and interpretation of the results, and wrote he clinical study report')	≥ 20 (NR) [NR]	Non-valvular ECG diagnosed $CHADS_2 \geq 1$	536	Edoxaban: 1. 30mg daily 2. 45mg daily 3. 60mg daily Warfarin: 4. INR 2-3	3	83 (≥ 70 years) 73 (< 70 years)	**Efficacy:** stroke or SE **Safety:** all bleeding, major bleeding, CRNM bleeding, composite CRB	3
AF-EDOX-VKA-MULTI (272) (Multi-centre) [North America, Chile, Europe and Russia]	Phase II Daiichi Sanyo Co. (not clear)	18-85 (65.1) [62.1]	Persistent non-valvular ECG diagnosed $CHADS_2 \leq 2$	1,146	Edoxaban: 1. 30mg daily 2. 65mg daily 3. 30mg bd 4. 60mg bd Warfarin: 5. INR 2-3	3	49.7	**Efficacy:** stroke or SE, MI, hospital admission **Safety:** all bleeding, major bleeding, minor bleeding, CRNM bleeding, composite CRB, death (cardiovascular)	3

AF-VKA-ASA-CHINA (273) (Multi-centre) [China]	Phase III 10th National Five-Year Project of China (not applicable)	50-80 (NR) [NR]	Non-valvular Diagnosis based on medical history, ECG and / or Holter recordings	690	Warfarin: 1. INR 2.1-2.5 2. INR 1.6-2.0 Antiplatelet: 3. Aspirin 200mg daily	24 (mean 15)	NR	**Efficacy:** all stroke, ischaemic stroke, haemorrhagic stroke, TIA **Safety:** major bleeding, minor bleeding, CRNM bleeding, composite CRB, death (cardiovascular)	24
ARISTOTLE (274-284) (Multi-centre) [North and South America, Europe, Russia, Israel, Australia, Asia and South Africa]	Phase III Bristol-Myers Squibb and Pfizer (The trial was designed in conjunction with the Sponsors and 'the primary analyses were performed both at BMS and the Duke Clinical Research Institute')	≥ 18 (median 70) [64.7]	Non-valvular or flutter ECG diagnosed	18,201	Apixaban: 1. 5mg bd (2.5mg bd in participants with more than one of ≥ 80 years, ≤ 60kg, serum creatinine level of ≥ 1.5mg/dL) Warfarin: 2. INR 2-3	21.6 (median)	62.2	**Efficacy:** all stroke, ischaemic stroke, haemorrhagic stroke, stroke or SE, MI **Safety:** all bleeding, major bleeding, composite CRB, intracranial bleeding, death (all causes)	21.6 (median for intracranial bleeding)
ARISTOTLE-J (285) (Multi-centre) [Japan]	Phase II Pfizer and Bristol-Myers Squibb (not clear)	≥ 20 (70.3) [82.9]	Non-valvular Diagnosis based on ECG, Holter recroding or intracardiac electrogram	222	Apixaban: 1. 2.5mg bd 2. 5mg bd Warfarin: 3. INR 2-3 (INR 2-2.6 in ≥ 70 years)	3	60	**Efficacy:** stroke or SE, ischaemic stroke, TIA **Safety:** all bleeding, major bleeding, minor bleeding, CRNM bleeding, composite CRB, death (all causes)	3

AVERROES (270, 275, 276, 286) (Multi-centre) [North and South America, Europe, Russia, Israel, Australia, Asia and South Africa]	Phase III Bristol-Myers Squibb and Pfizer (The Sponsors was involved in the design, data collection and analysis)	≥ 50 (70) [58.5]	Non-valvular ECG diagnosed	5,599	Apixaban: 1. 5mg bd (2.5mg bd if > 80 years or ≤ 60kg / renal status) Antiplatelet (Aspirin): 2. 81-324mg daily	13.1 (median)	NR	**Efficacy:** all stroke, stroke or SE, ischaemic stroke, haemorrhagic stroke, MI **Safety:** major bleeding, minor bleeding, CRNM bleeding, intracranial bleeding, fatal bleeding, death (cardiovascular), death (all causes)	13.1 (median)
BAFTA (286) (Multi-centre) [UK]	Phase III The Medical Research Council UK and supported by MidRec and the Primary Care Research Trust (The Sponsors had no direct role in study design, data collection, analysis or interpretation, writing the report, or decision to submit for publication)	≥ 75 (81.5) [54.6]	Non-valvular or atrial flutter ECG diagnosed	973	Antiplatelet (Aspirin): 1. 75mg daily Warfarin: 2. INR 2-3	32.4 (mean)	67	**Efficacy:** all stroke, MI **Safety:** major bleeding, death (all causes)	32.4 (mean)
Chinese ATAFS (287) (Multi-centre) [China]	Phase III (not disclosed)	40-80 (63.3) [59.7]	Non-valvular	704	Antiplatelet (Aspirin): 1. 150-160mg daily Warfarin: 2. INR 2-3 (INR 1.6-2.5 in > 75 years)	NR	NR	**Efficacy:** all stroke **Safety:** death (all causes)	19 (median) [range 2-24]

ENGAGE AF-TIMI-48 (288, 289) (Multi-centre) [North and South America, Europe, Russia, Israel, Australia, Asia and South Africa]	Phase III Daichii Sankyo Pharma Development (not clear)	≥ 21 (NR) [61.9]	Non-valvular ECG diagnosed $CHADS_2$ ≥ 2	21,105	Edoxaban: 1. 30mg daily 2. 60mg daily (half dose where creatinine clearance is 30-50 ml/min, ≤ 60kg, or concomitant use of verapamil or quinidine or dronedarone) Warfarin: 3. INR 2-3	29.8 (median)	64.9	**Efficacy:** all stroke, ischaemic stroke, haemorrhagic stroke, fatal stroke, stroke or SE, MI **Safety:** major bleeding, minor bleeding, fatal bleeding, intracranial bleeding, CRNM bleeding, composite CRB, death (cardiovascular), death (all causes)	29.8 (median)
EXPLORE-Xa (279) (Multi-centre) [USA, Canada and Germany]	Phase II Portolo Pharmaceuticals, South San Francisco (not stated)	≥ 18 (73) [66.5]	New or existing non-valvular or atrial flutter Diagnosis based on ECG, Holter recroding, rhythm strip, pacemaker or other intracardiac reading	508	Betrixaban: 1. 40mg daily 2. 60mg daily 3. 80mg daily Warfarin: 4. INR 2-3	4.9 (mean)	63.4	**Efficacy:** all stroke **Safety:** all bleeding, major bleeding, minor bleeding, CRNM bleeding, composite CRB, death (all causes)	4.9 (mean)

J-ROCKET AF (290) (Multi-centre) [Japan]	Phase III Bayer Yakuhin Ltd (The funder was 'responsible for trial design and study data collection')	≥ 20 (71.1) [80.6]	Non-valvular ECG diagnosed	1,280	Rivaroxaban: 1. 15mg daily (10mg daily if creatinine clearance 30-49 ml/min) Warfarin: 2. INR 2-3 (INR 1.6-2.6 in ≥ 70 years)	30	65	**Efficacy:** all stroke, ischaemic stroke, haemorrhagic stroke, stroke or SE, MI **Safety:** composite CRB, death (cardiovascular), death (all causes)	30
PATAF (265) (Multi-centre) [Netherlands]	Phase III Prevention fund (grant 002817010), Zorgonder-Zoek, Netherlands (not stated)	≥ 60 (74.8) [44.9]	Chronic or intermittent ECG diagnosed	729	Warfarin: 1. INR < 2 2. INR 2.5-3.5 (some patients received other coumarins – pheonprocoumon or acenocoumarol) Antiplatelet (Aspirin): 3. 150mg daily	32.4 (mean)	NR	**Efficacy:** all stroke, ischaemic stroke, arterial event **Safety:** death (cardiovascular), death (all causes)	32.4 (mean)

PETRO (291) (Multi-centre) [USA, Denmark, Netherlands and Sweden]	Phase II Boehinger Ingelheim (the funder 'was responsible for the statistical analysis conducted according to a prospectively designed plan approved by the steering committee')	≥ 18 (69.5) [81.9]	Permanent, persistent and paroxysmal non-valvular with coronary arterial disease Diagnosis not explained	502	Dabigatran: 1. 50mg bd 2. 50mg (+ aspirin 81mg) bd 3. 50mg (+ aspirin 325mg) bd 4. 150mg bd 5. 150mg (+ aspirin 81mg) bd 6. 150mg (+ aspirin 325mg) bd 7. 300mg bd 8. 300mg (+ aspirin 81mg) bd 9. 300mg (+ aspirin 325mg) bd Warfarin: 10. INR 2-3	3	57.2	**Efficacy:** stroke or SE **Safety:** all bleeding, major bleeding, composite CRB	3

RE-LY (292, 293) (Multi-centre) [North and South America, Europe, Russia, Israel, Australia, Asia and South Africa]	Phase III Boehinger Ingelheim (the Sponsor contributed to the design, conduct, and reporting of the study)	≥ 18 (71) [63.6]	Non-valvular ECG diagnosed Mean $CHADS_2$ = 2.1	18,113	Dabigatran: 1. 110mg bd 2. 150mg bd Warfarin: 10. INR 2-3	24 (mean)	64	**Efficacy:** stroke or SE, ischaemic stroke, haemorrhagic stroke, MI, PE, hospital admission **Safety:** major bleeding, minor bleeding, intracranial bleeding, extracranial minor bleeding, death (cardiovascular), death (all causes)	24 (mean)
ROCKET-AF (294-297) (Multi-centre) [North and South America, Europe, Russia, Israel, Australia, New Zealand, Asia and South Africa]	Phase III Johnson & Johnson, Bayer (the Sponsor was not involved in the coordination of the trial, data management or analyses)	≥ 18 (median 73) [60.3]	Non-valvular ECG diagnosed Mean $CHADS_2$ ≥ 2	14,264	Rivaroxaban: 1. 20mg daily (15mg daily if creatinine clearance 30-49 ml/min) Warfarin: 2. INR 2-3	19.4 (median)	55	**Efficacy:** all stroke, stroke or SE, MI **Safety:** major bleeding, CRNM bleeding, composite CRB, fatal bleeding, intracranial bleeding, death (all causes)	19.4 (median)

SPAF II (298) (Multi-centre) [USA]	Phase III The Division of Stroke an Trauma, National Institute of Neurological Disorders and Stroke (not clear)	Not clear (NR) [NR]	Non-valvular	1,100	Warfarin: 1. INR 2-4.5 (in < 75 years) 2. INR 2-4.5 (in > 75 years) Antiplatelet (Aspirin): 3. 325mg daily (in < 75 years) 4. 325mg daily (in > 75 years)	37.2 (mean for < 75 years) 24 (mean for > 75 years)	NR	**Efficacy:** stroke or SE, ischaemic stroke, MI, TIA **Safety:** intracranial bleeding, death (all causes)	27.6 (mean)
WASPO (299) (Multi-centre) [UK]	Phase III Not declared	>80 and <90 (median 83) [47]	Permanent non-valvular	75	Warfarin: 1. INR 2-3 Antiplatelet (Aspirin): 2. 300mg daily	12	69.2	**Efficacy:** all stroke, TIA **Safety:** death (all causes)	12

7.8.2 Time in Therapeutic Range (TTR) for warfarin

Table 35 indicates what the comparator intervention, target INR and (where reported) mean TTR was for the studies that included a warfarin intervention arm. Sixteen (73%) of these studies reported mean TTR, which varied substantially (from 45.1% to 83%) between studies.

Table 35: Mean TTR for warfarin in stroke prevention in AF

Study	Interventions that were compared with warfarin	Warfarin INR	Mean TTR (INR), %
ACTIVE W (265)	Antiplatelet (clopidogrel 75mg + (aspirin 75-100mg) once daily	2-3	63.8
AFASAK (266)	Aspirin 75mg once daily, Placebo once daily	2-3	73
AFASAK II (267)	Aspirin 300mg once daily	2-3	73
AF-ASA-VKA-CHINA (268)	Aspirin 100mg once daily	1.6-2.5	NR
AF-DABIG-VKA-JAPAN (269)	Dabigatran 110mg or 150mg twice daily	2-3 (≥1.6 to ≤2.6 in ≥ 70 years)	NR
AF-EDOX-VKA-ASIA (270)	Edoxaban 30mg or 60mg once daily	2-3	45.1
AF-EDOX-VKA-JAPAN (271)	Edoxaban 30mg, 45mg or 60mg once daily	2-3 (1.6 to 2.6 in ≥ 70 years)	83 (≥ 70) 73 (< 70)
AF-EDOX-VKA-MULTI (272)	Edoxaban 30mg or 60mg once daily, Edoxaban 30mg or 60mg twice daily	2-3	49.7
AF-VKA-ASA-CHINA (273)	Aspirin 200mg once daily	2.1-2.5	NR
ARISTOTLE (274-284)	Apixaban 5mg twice daily	2-3	62.2
ARISTOTLE-J (285)	Apixaban 2.5mg or 5mg twice daily	2-3 (2 to 2.6 in ≥ 70 years)	60
BAFTA (286)	Aspirin 75mg once daily	2-3	67
Chinese ATAFS (287)	Aspirin 150mg to 160mg once daily	2-3 (1.6 to 2.6 in > 75 years)	NR
ENGAGE AF-TIMI-48 (288, 289)	Edoxaban 30mg or 60mg once daily	2-3	64.9
EXPLORE-Xa (279)	Betrixaban 40mg or 60mg or 80mg once daily	2-3	63.4
J-ROCKET AF (290)	Rivaroxaban 15mg once daily	2-3 (1.6 to 2.6 in ≥ 70 years)	65
PATAF (265)	Aspirin 150mg once daily	2.5-3.5	NR

Study	Interventions that were compared with warfarin	Warfarin INR	Mean TTR (INR), %
PETRO (291)	Dabigatran 50mg, or 50mg + aspirin 81mg, or 50mg + aspirin 325mg, or 150mg, or 150mg + aspirin 81mg, or 150mg + aspirin 325mg, or 300mg, or 300mg + aspirin 81mg, or 300mg + aspirin 325mg (all twice daily)	2-3	57.2
RE-LY (292, 293)	Dabigatran 110mg or 150mg twice daily	2-3	64
ROCKET-AF (294-297)	Rivaroxaban 20mg once daily	2-3	55
SPAF II (298)	Aspirin 325mg once daily	2-4.5	NR
WASPO (299)	Aspirin 300mg once daily	2-3	69.2

7.8.3 Risk of bias

A risk of bias assessment was undertaken for each included study using the Cochrane assessment tool against each domain (see Table 36). Each assessment was made by me and verified by Mr Okoli.

The assessments ranged from low to high risk of bias; however it was difficult to judge some studies as a result of inaccessibility of study protocols. For most of the outcomes assessed in the studies, all randomised patients were either accounted for in the analysis, or in some cases a small number of patients were unaccounted for with reasons judged likely to be unrelated to the outcome.

The majority of the studies were judged to be at low risk of bias for allocation concealment and incomplete outcome data. The majority of the studies were judged to be at a low or unclear risk of bias for sequence generation. Randomisation sequence across the low-risk studies was predominantly computerised. Most studies were also judged to be of low risk of bias for blinding of outcome assessment, with three studies judged to be at high risk of bias in this domain. Fourteen studies were judged to be at high risk of bias for blinding of participants and personnel, mainly because they were open-label. Where studies were blinded for different dose groups of a novel anticoagulant, but not in the comparison of these to warfarin, we assigned a high risk of bias because the principal contribution of the study to our analyses would be the comparison of warfarin with the licensed dose of the anticoagulant. Risk-of-bias judgements for studies contributing to analyses of each outcome are presented graphically in the sections that follow.

Table 36: Risk of bias assessment for NOAC studies included within the analysis of stroke prevention in AF

Study	Sequence generation (SG)	Allocation concealment (AC)	Blinding of participants and personnel (BPP)	Blinding of outcome assessments (BOA)	Incomplete outcome data (IOD)	Selective Reporting (SR)
ACTIVE W (265)	L 'Patients were randomised by an automated central interactive voice-response system, in a 1 : 1 ratio, to receive clopidogrel plus aspirin or oral anticoagulation therapy'	L By means of a central, interactive, voice response system	H Treatment was open, with blinded adjudication of outcomes	L 'All major outcomes were adjudicated by a blinded committee and all strokes were adjudicated by neurologists'	L All patients were included in the analyses	U Study protocol not found
AFASAK (266)	L 'The patients were randomised to receive warfarin, aspirin 75 mg od, or placebo. They received consecutive numbers, which corresponded to numbered packages containing the study medication, the order of which was determined by computer generated randomisation'	L 'They received consecutive numbers, which corresponded to numbered packages containing the study medication, the order of which was determined by computer-generated randomisation'	H 'Warfarin was given openly, but the aspirin and placebo arms were double-blind. The warfarin tablets looked different from the aspirin and placebo tablets, which were indistinguishable'	U No information on blinding of outcome assessors	L All patients were included in the analyses	U Study protocol not found
AFASAK II (267)	L 'According to a computer-generated sequence, eligible participants were assigned to daily treatment'	U No information on whether or note and how treatment was concealed	H This was an open-label study	L 'All end-points were evaluated by an end-point committee unaware of treatment status. The committee consisted of two neurologists and two cardiologists'	L All patients were included in the analyses	U Study protocol not found

AF-ASA-VKA-CHINA (268)	U 'A total of 110 patients met the inclusion criteria and were randomly divided into warfarin study and aspirin control groups'	H No information and no indication of treatment allocation concealment	H No details, but monitoring of INR implies the study was open-label	H No information, and no indication of blinding outcome assessors	L Small numbers of missing data in the two randomised arms and the number missing in each arm seem to be balanced; also with comparable reasons for the missing data. It is unlikely that missing data are related to the outcome	U Study protocol not found
AF-DABIG-VKA-JAPAN (269)	U Available information is from a preliminary report and not a published article. Therefore no information to enable judgement	U Available information is from a preliminary report and not a published article. Therefore no information to enable judgement	U Available information is from a preliminary report and not a published article. Therefore no information to enable judgement	U Available information is from a preliminary report and not a published article. Therefore no information to enable judgement	U Available information is from a preliminary report and not a published article. Therefore no information to enable judgement	U Available information is from a preliminary report and not a published article. Therefore no information to enable judgement
AF-EDOX-VKA-ASIA (270)	L 'Via Fisher Automated Clinical Trials System (FACTS)'	L 'Block randomisation was done by FACTS; Cenduit produced the randomisation schedule; which was kept confidential until the end of the study'	H 'The investigator, patients and sponsor were blinded to the dose of edoxaban, but not to the identity of edoxaban and warfarin'	L 'The independent CEC, which was blinded to study treatments, adjudicated all bleeding events and thromboembolic events (stroke, systemic embolic event, MI) during the study'	L Only one person with missing outcome data	L All outcomes are reported as per protocol
AF-EDOX-VKA-JAPAN (271)	L 'Treatment was assigned using the biased coin method'	U 'Patients were randomised using the specifications of dynamic allocation procedures'	H 'This was a multicentre, randomised, dose-ranging study of edoxaban (double-blind to dose) and open-label warfarin'	U 'Secondary endpoints consisted of thromboembolic events including stroke assessed by an independent Event Assessment Committee'	L Some missing data with reasons although the number of missing data are quite minimal and unlikely related to the outcome	L All outcomes are reported as per protocol

AF-EDOX-VKA-MULTI (272)	L 'The randomisation schedule was generated by an independent biostatistician who was not part of the study team. Using a central, interactive, automated telephone system'	L 'Using a central, interactive, automated telephone system, eligible patients who provided written informed consent were randomly allocated'	H 'The study was double blind with respect to edoxaban dose, but open-label for randomisation between edoxaban and warfarin'	U For efficacy outcomes: 'Stroke confirmed by CT or sutopsy; TIA confirmed by a neurologist' L For safety outcomes: 'Suspected bleeding events were assessed by an independent blinded adjudication committee'	L Very minimal missing data in three arms (1 patient); otherwise all patients were accounted for in the analyses	L All outcomes are reported as per protocol
AF-VKA-ASA-CHINA (273)	L 'Stratified block randomisation'	U Not enough information: 'After giving signed informed consent, patients who met the inclusion criteria were enrolled and randomly allocated to one of three study groups according to a stratified block randomisation'	U 'In the warfarin groups, an initial dose of 1–3mg/day of warfarin was prescribed after the baseline INR values were measured'. 'In the aspirin group, a fixed dose of 200 mg/day of aspirin was used'	U Not clearly described 'Medical records from all potential events were further reviewed by a five-physician clinical outcomes committee'	U A total of 96 patients withdrew from the study after randomisation but the remaining 690 patients were all included in the analysis	U Study protocol not found

ARISTOTLE (274-284)	L 'Randomisation was stratified according to whether patients had received warfarin previously and according to clinical site'	U 'An algorithm was provided to manage temporary discontinuations of the study drug around the time of interventional procedures while maintaining concealment of the group assignments'	L 'An algorithm was provided to manage temporary discontinuations of the study drug around the time of interventional procedures while maintaining concealment of the group assignments'	L 'The primary and secondary efficacy and safety outcomes were adjudicated on the basis of prespecified criteria by a clinical events committee whose members were not aware of study-group assignments'	L For efficacy outcomes: No missing outcome data U For bleeding outcomes: Some missing outcome data with reason which appear to be similar in the groups. However, it is not clear whether or not the reasons for the missing outcome data are related to the outcome	L All outcomes are reported as per protocol
ARISTOTLE-J (285)	U 'Patients were randomised in a 1 : 1 : 1. The randomisation assignment method (Pocock et al.) incorporated trial site and warfarin status (experienced or naive) as factors'	U Not enough information. 'On the first day of study drug dosing (week 0), patients were randomised in a 1 : 1 : 1 fashion'	H 'This was a randomised, partially blinded study comparing high double-blinded doses of apixaban with open-label warfarin'	L 'An independent blinded end-point committee adjudicated all reported bleeding and efficacy events'	L Few outcome missing data with reasons. Reasons almost balance out across groups and it is unlikely that the reasons are related to the outcome	L All outcomes are reported as per protocol

AVERROES (270, 275, 276, 286)	L 'Randomisation was performed with the use of a 24-hour central, computerised, automated voice-response system'	L 'Randomisation was performed with the use of a 24-hour central, computerised, automated voice-response system'	L 'In keeping with the double-dummy design, patients who were assigned to receive apixaban also received an aspirin placebo, and those assigned to receive aspirin also received an apixaban placebo'	L 'All outcomes were adjudicated by a committee whose members were unaware of the treatment assignments. Cases of stroke and ICH were adjudicated by neurologists'	L All patients were included in the analyses	L All outcomes are reported as per protocol
BAFTA (286)	L 'Within each stratum, randomly permuted blocks of eight were generated to produce allocation tables'	L 'Primary care physicians telephoned for the treatment allocation when they had an eligible patient'	H 'BAFTA was a prospective randomised open-label trial'	L 'Clinical details on possible primary events were sent to two independent neurologists who were blind to treatment allocation'	L All patients were included in the analyses	L All outcomes are reported as per protocol
Chinese ATAFS (287)	U 'The randomised study of efficacy and safety of antithrombotic therapy in non-valvular AF: warfarin compared with aspirin'	U No information on whether or not and how treatment allocation was concealed	U No information on whether or not participants and personnel were blinded to treatment	U No information on blinding of outcome assessors	L All patients were included in the analyses	U Study protocol not found

ENGAGE AF-TIMI-48 (288, 289)	L 'Randomisation was performed with the use of a central, 24-hour, interactive, computerised response system'	L 'Randomisation was performed with the use of a central, 24-hour, interactive, computerised response system'	L 'Each patient received two sets of study drugs: either active edoxaban and a placebo matching warfarin, or a placebo matching edoxaban and active warfarin'	L 'An independent clinical end-point committee, whose members were unaware of the study assignment, adjudicated all deaths and suspected cerebrovascular events, systemic embolic events, MIs, bleeding events, and hepatic events'	L All patients were included in the analyses	L All outcomes are reported as per protocol
EXPLORE-Xa (279)	U 'Patients were randomly assigned (1 : 1 : 1 : 1 allocation) A dynamic randomisation was used to assign and balance patients by country, concurrent aspirin use, and antecedent warfarin'	U Not enough information. 'Patients were randomly assigned (1 : 1 : 1 : 1 allocation)'	H 'Assignment to betrixaban or warfarin was not blinded, but the betrixaban dose was double-blinded'	L 'An independent adjudicator, blinded to treatment groups, adjudicated all major bleeds, CRNM bleeds, strokes, MI, other SE and deaths'	L All patients were included in the analyses	L All outcomes are reported as per protocol
J-ROCKET AF (290)	L No details provided but assumed to follow robust design of the ROCKET AF study	L No details provided but assumed to follow robust design of the ROCKET AF study	L 'As part of the doubledummy design, patients in each group also received a tablet of either titrated warfarin placebo or rivaroxaban placebo, respectively, to preserve the treatment blind'	U 'An independent clinical end-point committee adjudicated all suspected strokes, SEs, MIs, deaths, and bleeding events contributing to the prespecified end-points'	L Very few missing data. Unlikely to influence the true outcome	L All outcomes are reported as per protocol

PATAF (265)	L Randomisation was computer generated	L 'Patients eligible for standard anticoagulation were randomly assigned (centrally, by telephone)'	U 'Patients were single blinded for the two intensities of anticoagulant'	L 'End-point ascertainments were blinded for treatment. Events were independently reviewed by two members of the (neurological, cardiological, vascular, ophthalmological and internal medicine) event committees (or three, in case of disagreement'	U Some missing data and although with similar reasons across groups, the missing numbers in the groups are not balanced	U Study protocol not found
PETRO (291)	U 'The PETRO study was a randomised trial of patients with AF at high risk for thromboembolic events'	U Not enough information 'Randomisation was stratified in the ratio 6 : 9 : 9 : 4 (50-, 150- and 300-mg dabigatran, and warfarin, respectively)'	H 'The trial was doubleblind with respect to dabigatran dose but open-label for concomitant aspirin treatment'	U For efficacy outcome: No information but the outcomes may have been blinded L For bleeding outcome: 'An independent adjudication committee blinded to treatment evaluated all bleeding events'	L All patients were included in the analyses	L All outcomes are reported as per protocol

RE-LY (292, 293)	L 'After providing written informed consent, all trial participants were randomly assigned to receive one of two doses of dabigatran, or to receive warfarin, by means of a central, interactive, automated telephone system'	L By means of a central, interactive, automated telephone system	H 'RE-LY was a randomised trial designed to compare two fixed doses of dabigatran, each administered in a blinded manner, with open-label use of warfarin'	L 'Each primary and secondary outcome event was adjudicated by two independent investigators who were unaware of the treatment assignments'	L All patients were included in the analyses	L All outcomes are reported as per protocol
ROCKET-AF (294-297)	L 'Randomisation was performed with the use of a central 24-hour, computerised, automated voice-response system'	L 'Randomisation was performed with the use of a central 24-hour, computerised, automated voice-response system'	L 'Patients were randomly assigned to receive fixed-dose rivaroxaban or adjusted dose warfarin. Patients in each group also received a placebo tablet in order to maintain blinding'	U 'An independent clinical end-point committee applied protocol definitions to adjudicate all suspected cases of stroke, SE, MI, death and bleeding events that contributed to the pre-specified end-points'	L Very few missing outcome data but with reasons which appear to balance across groups. Unlikely that the missing data are related to the true outcome	L All outcomes are reported as per protocol

SPAF II (298)	L Randomisation was done separately at each clinical site by computer	U Not enough information: 'The randomisation sequence could not be pre-reviewed'	H Both patient and investigator were aware of therapy assignment	L For neurological efficacy outcomes: 'All suspected neurological events were evaluated by an on-site study neurologist and verified by an events committee; evaluation was based on review of original medical records, from which information about therapy assignment had been removed' H For safety outcomes: No details on blinding of outcome assessment	L All patients were included in the analyses	U Study protocol not found
WASPO (299)	L 'Randomisation was prepared from a computer-generated random numbers programme'	L 'Randomisation was performed by opening sealed envelopes in numbered sequence'	H This was an open-label study	H No information and no indication of blinding	L All patients were included in the analyses	U Study protocol not found

(H) high risk; (L) low risk; (U) unclear risk
Quotations from the publication articles are denoted by inverted commas

7.8.4 Results on clinical effectiveness and safety

The 27 different interventions considered in the 23 trials are listed in Table 37 and Table 38, which show the number of patients analysed and the number of outcome events for each outcome reported in each trial. NMAs were performed for seven outcomes:

- stroke or SE
- ischaemic stroke
- MI
- major bleeding
- CRB
- intracranial bleeding
- all-cause mortality

Arms that were considered not to provide any evidence of interest to inform health decisions in the UK were excluded from the analyses. Specifically, the warfarin arm with INR range 1.6–2 from the AF-VKA-ASA-CHINA trial (273), the warfarin arm with INR range of < 2 from PATAF (265), the placebo arm from AFASAK (266), and the two warfarin arms with a fixed daily dose from AFASAK II (267) were excluded.

Two independent nodes were defined for warfarin interventions, labelled as ‘warfarin (INR 2–3)’ and ‘warfarin (INR 3–4)’, respectively. The first of these formed the reference treatment across all networks in the AF review as this is the UK standard of care. The ‘warfarin (INR 2–3)’ node included the trials with a therapeutic INR range of 2–3 (e.g. ACTIVE W (300), AFASAK (266)), as well as some interventions with an INR range of 2.5–3.5 (AF-EDOX-VKA-ASIA (270) and PATAF (265)) or 2.0–4.5 (SPAF II (298)). In other trials the INR range for some patients in the warfarin arm was sub-therapeutic (< 2.0), so that the total INR range was 1.6–3.0. These interventions were excluded from the main analysis, but merged with the INR 2–3 node in a sensitivity analysis. As a consequence, there were three two-arm trials (J-ROCKET AF (290), Chinese ATAFS (287) and AF-ASA-VKA-CHINA (268)) that were included only in sensitivity analyses.

Two independent nodes were defined for antiplatelet interventions (‘aspirin’ or ‘aspirin plus clopidogrel’), using the cut-off point of 150

mg, with the understanding that daily doses above that were appropriate for stroke prevention in AF, whereas lower doses were appropriate for secondary prevention of cardiovascular events. The dose range considered in the AVERROES trial (273, 301-303) (81–324 mg od) was much wider than in any other trial, and we included this intervention in the lower-dose node (< 150 mg once daily) because patients from that study had received a low daily dose. As a sensitivity analysis, the AVERROES trial was excluded from the network. Finally, the main analysis used a binomial model, assuming equal follow-up times across arms within trials and ignoring some variations in how results were reported. A separate analysis was performed for all outcomes taking into account the differences in duration of follow-up within and between trials, and the differences in the definition of event used across trials (e.g. total number of events vs. first events only).

Due to the scale of this project the results are presented as follows for each of the seven outcomes.

- First, network plots are provided to illustrate the comparisons of interventions made in the different trials (see Figure 56 to Figure 64).
- Second, efficacy and safety outcomes reported from the randomised controlled trials for stroke prevention in AF (see section 7.8.4.1 to 7.8.4.7).
- Third, the risk-of-bias assessments are illustrated, specific to the outcome for each trial included in the network.
- Fourth, results tables are presented for each intervention compared with the reference treatment (warfarin with a target INR range of 2–3).
- Fifth, results tables are presented for pairwise comparisons among licensed doses of the NOACs. For both sets of results tables, posterior median ORs and 95% credible intervals from Bayesian fixed-effects analyses are shown, although the latter are referred to as CIs for convenience. In these tables results are presented separately for any available direct evidence, for any indirect comparisons that can be made (excluding the

direct evidence) and for the NMA (which combines the direct and the indirect evidence). Comparisons from the NMA with a ratio between interval limits of > 9 were considered 'imprecisely estimated' and are presented at the bottom of each table (calculation of indirect evidence was not undertaken for imprecisely estimated comparisons).

- Sixth, a summary of results across outcomes is provided, in the form of a 'rank-o-gram', which illustrates the probability that each treatment is best, second best, and so on, for each outcome (see section 7.8.5.1)
- Lastly, forest plots of all contributing data, with ORs calculated using standard frequentist methods are presented in section (see section 10.6; Appendix VI: Forest Plots (stroke prevention in AF).

Table 37: List of interventions (for the NMA) examined by included randomised controlled trials in stroke prevention in AF

NMA reference number	Intervention
1	Warfarin (INR 2-3)
2	Warfarin (INR 1.6-3.0)
3	Warfarin (INR 3-4) once daily
4	Antiplatelet (<150mg once daily)
5	Antiplatelet (≥150mg once daily)
6	Dabigatran (50mg twice daily) + aspirin (81mg twice daily)
7	Dabigatran (50mg twice daily) + aspirin (325mg twice daily)
8	Dabigatran (150mg twice daily) + aspirin (81mg twice daily)
9	Dabigatran (150mg twice daily) + aspirin (325mg twice daily)
10	Dabigatran (300mg twice daily) + aspirin (81mg twice daily)
11	Dabigatran (300mg twice daily) + aspirin (325mg twice daily)
12	Apixaban (2.5mg twice daily)
13	Apixaban (5mg twice daily)
14	Dabigatran (50mg twice daily)
15	Dabigatran (110mg twice daily)
16	Dabigatran (150mg twice daily)
17	Dabigatran (300mg twice daily)
18	Betrixaban (40mg once daily)
19	Betrixaban (60mg once daily)
20	Betrixaban (80mg once daily)
21	Edoxaban (30mg once daily)
22	Edoxaban (45mg once daily)
23	Edoxaban (60mg once daily)
24	Edoxaban (30mg twice daily)
25	Edoxaban (60mg twice daily)
26	Rivaroxaban (15mg once daily)
27	Rivaroxaban (20mg once daily)

Table 38: Summary of list of outcomes reported and patient numbers from the included randomised controlled trials in stroke prevention in AF

	No. Studies	No. patients	No. events
Transient ischemic attack (TIA)	6	3767	60
All stroke	14	72149	2541
Stroke or systemic embolism	16	90427	3161
Ischemic stroke	8	50131	1406
Minor ischemic stroke	1	1007	3
Haemorrhagic stroke	7	49626	331
Fatal stroke	4	23762	285
Pulmonary embolism	1	18113	43
Bleeding	14	28921	6225
Major bleeding	12	43481	7627
Minor bleeding	19	84736	4265
Fatal bleeding	4	42069	214
Extra-cranial minor bleeding	1	18113	956
Intra-cranial bleeding	7	74265	722
MI	12	89632	1345
Hospital admission	2	19256	7211
Death (cardiovascular)	6	47628	2766
Death (all causes)	17	91867	6526
Arterial event	2	909	13
Clinically Relevant Non Major Bleed	10	38776	6045
Composite Clinically relevant bleeding	12	58126	9293

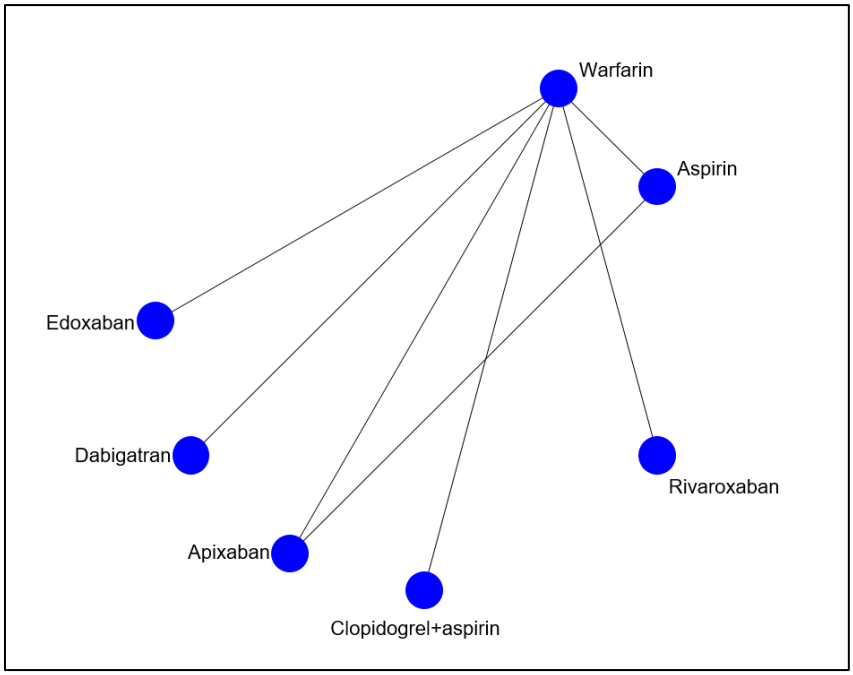


Figure 56: NOAC systematic review – Atrial fibrillation network plot (simple)

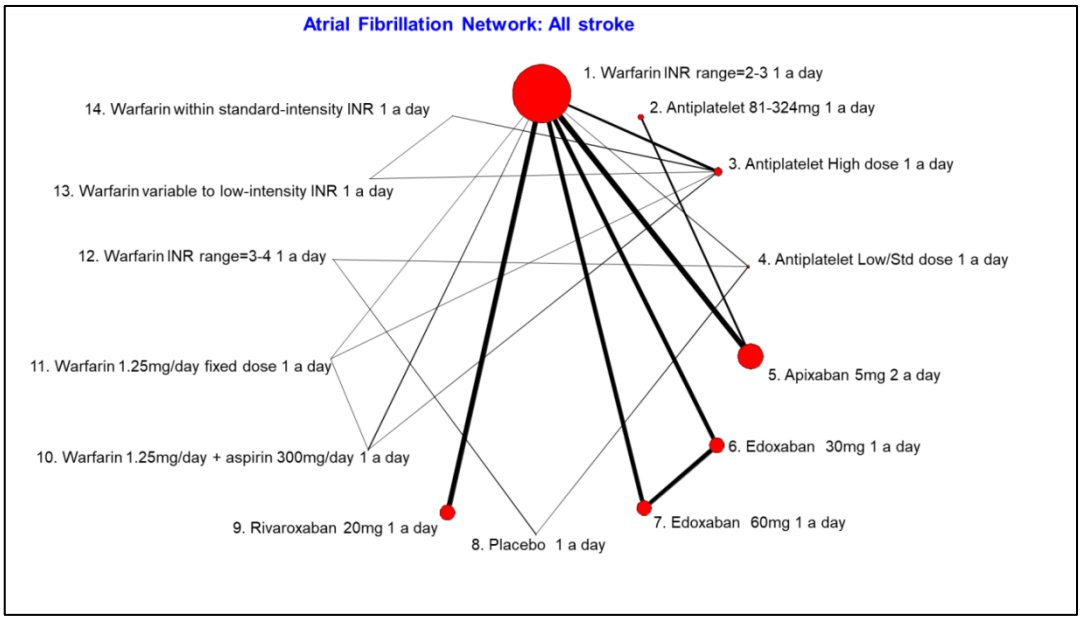


Figure 57: NOAC systematic review – Atrial fibrillation network plot (all stroke)

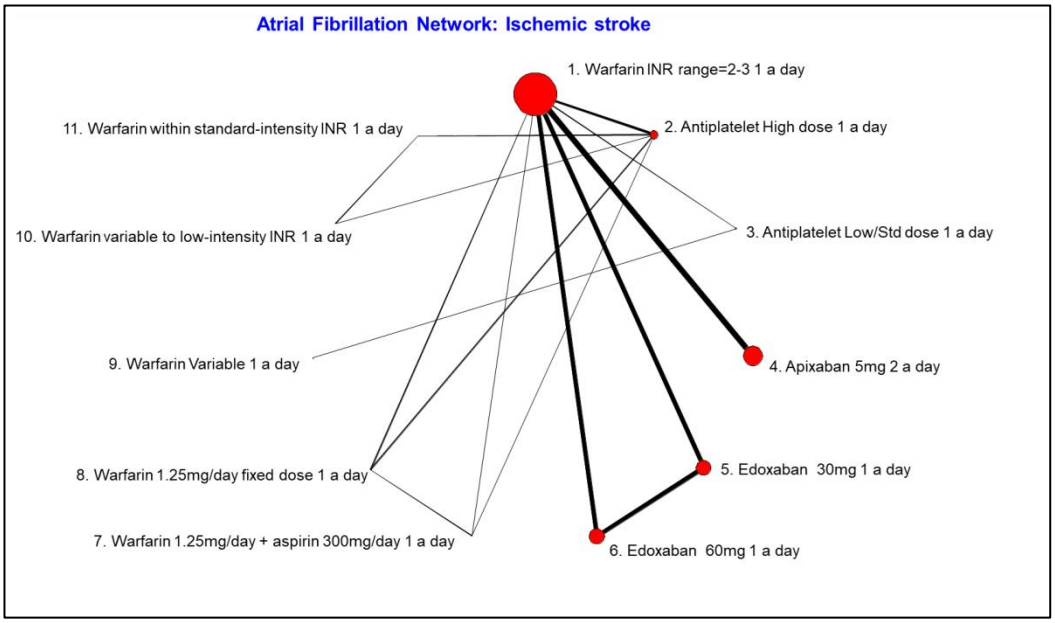


Figure 58: NOAC systematic review – Atrial fibrillation network plot (ischaemic stroke)

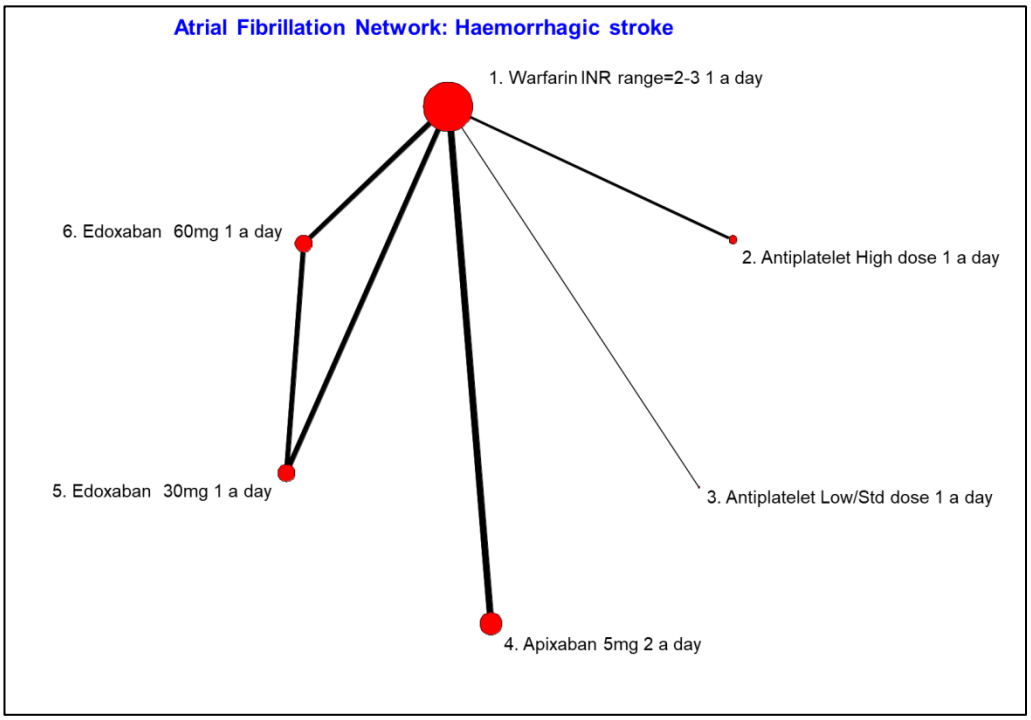


Figure 59: NOAC systematic review – Atrial fibrillation network plot (haemorrhagic stroke)

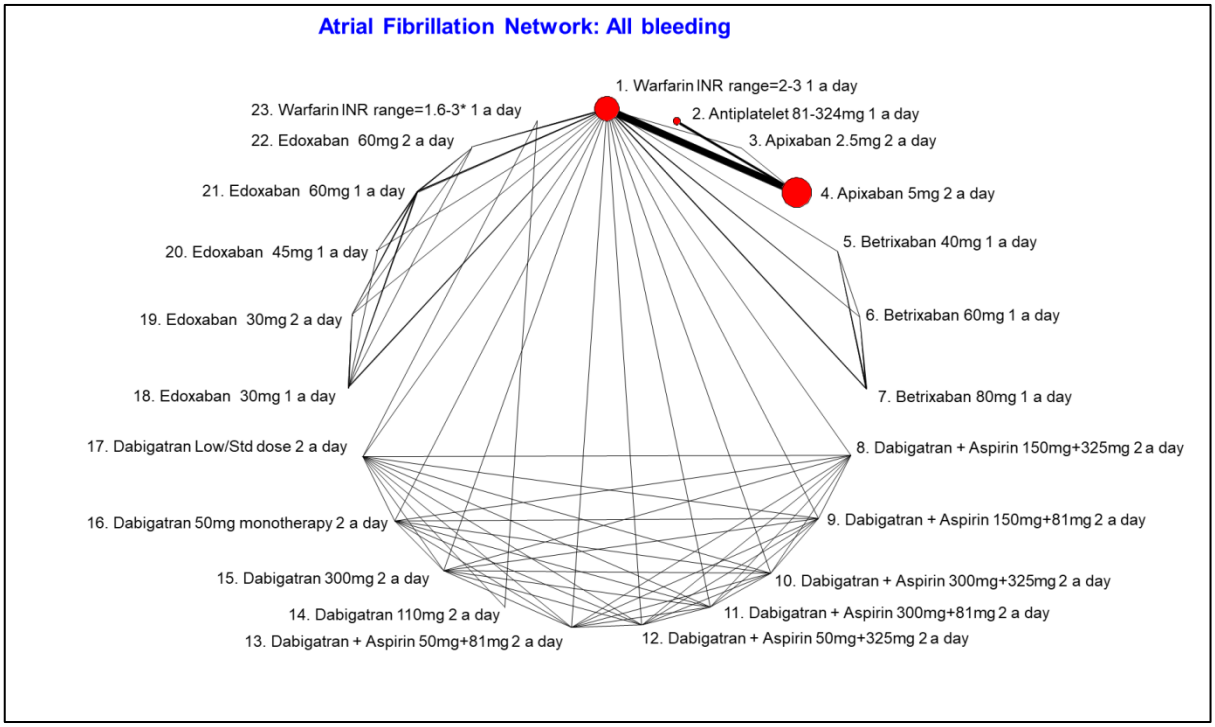


Figure 60: NOAC systematic review – Atrial fibrillation network plot (all bleeding)

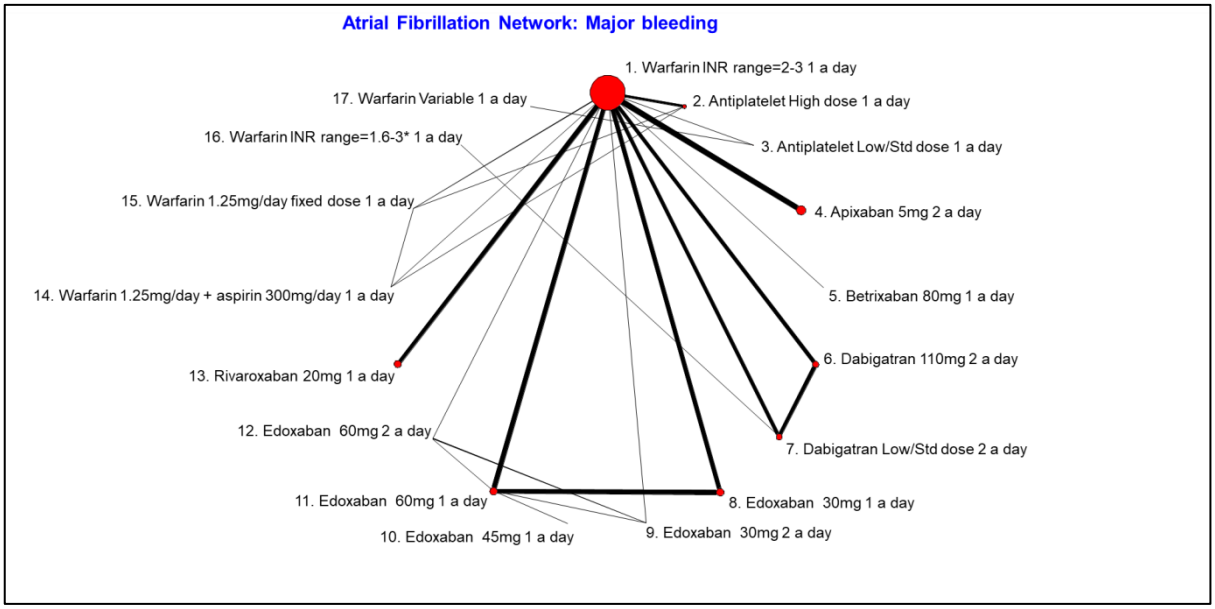


Figure 61: NOAC systematic review – Atrial fibrillation network plot (major bleeding)

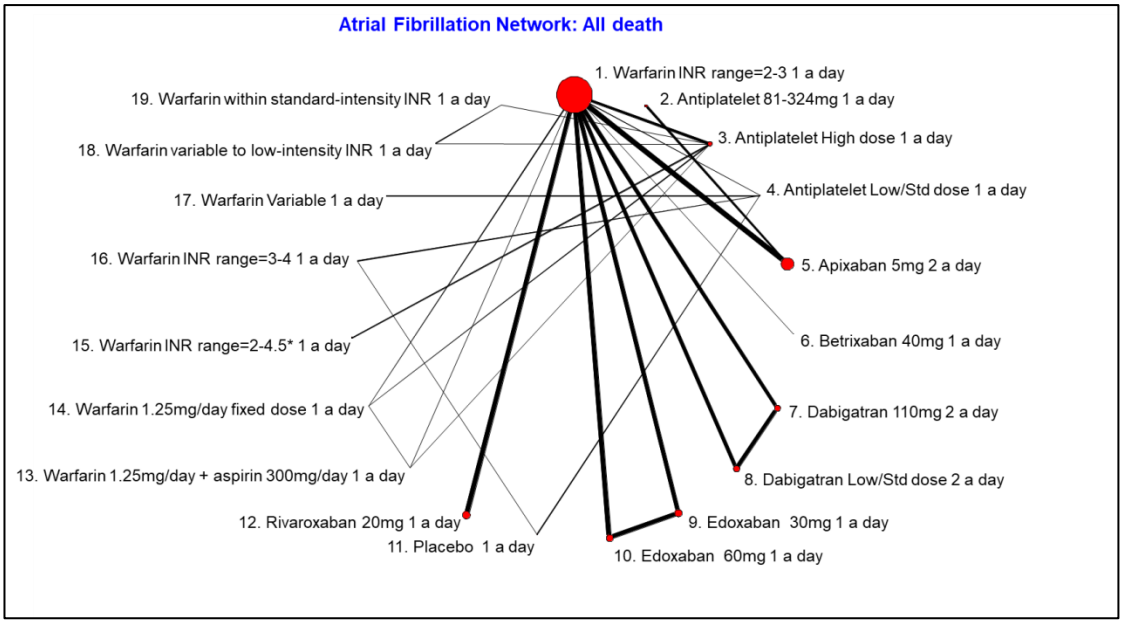


Figure 62: NOAC systematic review – Atrial fibrillation network plot (all death)

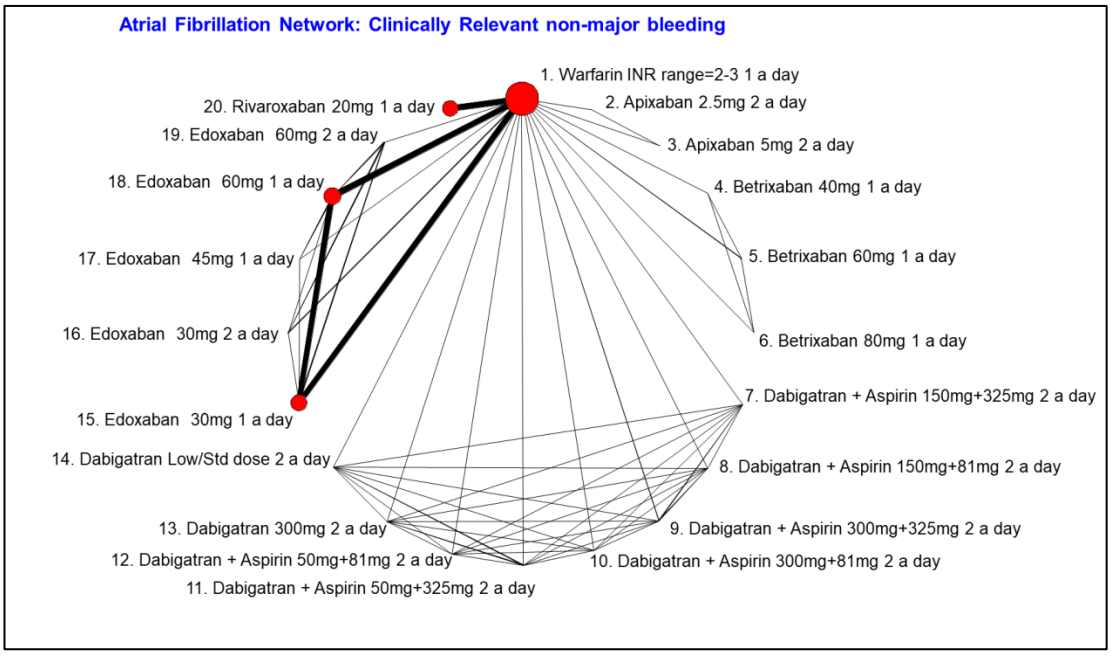


Figure 63: NOAC systematic review – Atrial fibrillation network plot (CRNM bleeding)

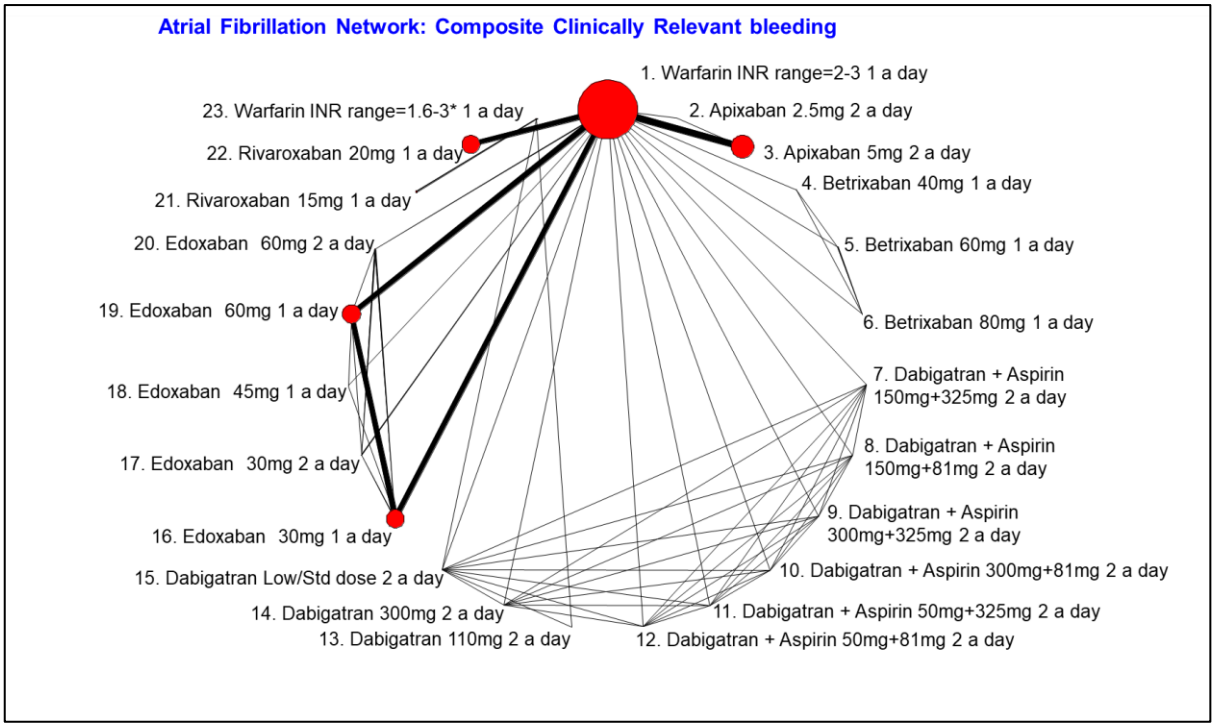


Figure 64: NOAC systematic review – Atrial fibrillation network plot (composite clinically relevant bleeding)

Table 39: Efficacy outcomes reported from the randomised controlled trials (stroke prevention in AF): number of events for each outcome at trial level

Study	Study size	TIA	All stroke	Stroke or SE	Ischaemic stroke	Minor ischaemic stoke	Major ischaemic stroke	Haemorrhagic stroke	Fatal stroke	PE	MI	Hospital admission
ACTIVE W (265)	**6,706**		159		132			20			59	
AF-ASA-VKA-CHINA (268)	**101**			18	14						5	
AF-DABIG-VKA-JAPAN (269)	**166**			1								
AF-EDOX-VKA-ASIA (270)	**234**			0								
AF-EDOX-VKA-JAPAN (271)	**519**			1								
AF-EDOX-VKA-MULTI (272)	**1,143**			11							5	12
AF-VKA-ASA-CHINA (273)	**440**	13	10		9			1				
AFASAK (266)	**671**	2	20			1			4			
AFASAK II (267)	**339**	3	19	22	8			2	2		8	
ARISTOTLE (274-284)	**18,140**		449	477	337			118			192	
ARISTOTLE-J (285)	**218**	1		3	1						0	

AVERROES (270, 275, 276, 286)	**5,599**		154	164	128			15			52	
BAFTA (286)	**973**		94								30	
Chinese ATAFS (287)	**704**		23									
ENGAGE AF-TIMI-48 (288, 289)	**21,026**		958	1,016	804			169	239		443	
EXPLORE-Xa (279)	**508**		2		2						0	
J-ROCKET AF (290)	**1,278**		31	33	24			7			4	
PATAF (265)	**272**		7		7	2	5				5	
PETRO (291)	**515**			2								
RE-LY (292, 293)	**18,113**			519	389			71		43	270	7,199
ROCKET-AF (294-297)	**14,236**		405	575	310						227	
SPAF II (298)	**1,100**	25		67	63						34	
WASPO (299)	**75**	1	0									

Table 40: Safety outcomes reported from the randomised controlled trials (stroke prevention in AF): number of events for each outcome at trial level

Study	Study size	All bleeding	Minor bleeding	Major bleeding	Fatal bleeding	Extracranial minor bleeding	Intracranial bleeding	Arterial bleeding	CRNM bleeding	CRB	Cardiovascular deaths	All-cause mortality
ACTIVE W (265)	**6,706**	1,199	1,049	194	18							317
AF-ASA-VKA-CHINA (268)	**101**	14	9	3	2							4
AF-DABIG-VKA-JAPAN (269)	**166**	45		3						14		
AF-EDOX-VKA-ASIA (270)	**234**	57	48	2					9	11		
AF-EDOX-VKA-JAPAN (271)	**519**	115		5					15	20		
AF-EDOX-VKA-MULTI (272)	**1,143**	114	52	13					49	62	8	
AF-VKA-ASA-CHINA (273)	**440**		25	8								11
AFASAK (266)	**671**	23										15
AFASAK II (267)	**339**		68	9			3					31
ARISTOTLE (274-284)	**18,140**	5,416		789			174			1,490		1,272
ARISTOTLE-J (285)	**218**	41	36	1					5	6		0

AVERROES (270, 275, 276, 286)	**5,599**		341	83	10		24		180	263	180	251
BAFTA (286)	**973**			50								215
Chinese ATAFS (287)	**704**											12
ENGAGE AF-TIMI-48 (288, 289)	**21,026**		1,851	1,196	112		234		3,579	4,456	1,668	2,349
EXPLORE-Xa (279)	**508**	118	109	8					12	18		2
J-ROCKET AF (290)	**1,278**						15			262	8	12
PATAF (265)	**272**							8			18	29
PETRO (291)	**515**	88		4						36		
RE-LY (292, 293)	**18,113**		5,284	1,162		956	150				880	1,371
ROCKET-AF (294-297)	**14,236**			781	82		139		2,336	2,924		458
SPAF II (298)	**1,100**						18					127
WASPO (299)	**75**		10	3								3

7.8.4.1 Stroke or Systemic Embolism

Sixteen studies reported the number of stroke or systemic embolism (SE) events, and the other seven trials reported the number of stroke events, so that the resulting network was based on data from all 23 trials, comparing a total of 26 interventions (see Figure 65).

There were 3,217 stroke or SE events in total. Twenty studies were included in the main analysis, with the remaining three included only in sensitivity analyses. The thicker lines joining interventions, which mainly correspond with comparisons between licensed doses of NOACs and warfarin (INR 2–3) represent the larger (mainly Phase III) trials. Similarly, the larger green circles represent the interventions to which the largest number of patients were randomised. Importantly, there were no direct comparisons between different NOACs, although there were numerous comparisons between different doses of the same NOAC in mainly Phase II trials, and some such comparisons in larger trials. Therefore, comparisons between the effects of different NOACs need to be inferred from the network (indirect evidence).

Table 41 shows risk-of-bias judgements for studies reporting stroke or SE. The studies were at mixed risks of bias: there were concerns about lack of blinding of participants for most trials, and about lack of allocation concealment and blinding of outcome assessment in some.

Table 42 shows comparisons of licensed doses with warfarin (INR 2–3), suggests that both low- and high-dose antiplatelet drugs increase the risk of stroke or SE compared with warfarin (INR 2–3). Among NOACs, there was some evidence that apixaban [5 mg bd (bd)], dabigatran (150 mg bd), edoxaban (60 mg od) and rivaroxaban (20 mg od) reduce the risk of stroke or SE compared with warfarin (INR 2–3). Most other comparisons were imprecisely estimated. Comparisons among licensed doses of NOACs were almost all based on indirect evidence (see Table 43). Among the

comparisons that were not classified as imprecisely estimated, there was some evidence that edoxaban (60 mg once daily) and rivaroxaban (20 mg once daily) increase the risk of stroke or SE compared with dabigatran (150 mg twice daily).

Results from a supplementary analysis taking into account the differences in duration of follow-up within and between trials, and the differences in the definition of event used across trials (e.g. total number of events vs. first events only), are presented in Table 44 and Table 45. These values are very similar to those for ORs.

As a post hoc sensitivity analysis, a fixed-effects meta-regression model was fitted using the mean TTR for warfarin patients (see Table 35) as a covariate and the mean log-odds ratio (log-OR) from each pairwise comparison (with warfarin as the reference category) as the response variable. There was little evidence of effect modification due to mean TTR (estimated coefficient 0.0021 with 95% CI −0.07 to 0.08 per 1% increase). The model fit indices were very similar with and without the covariate.

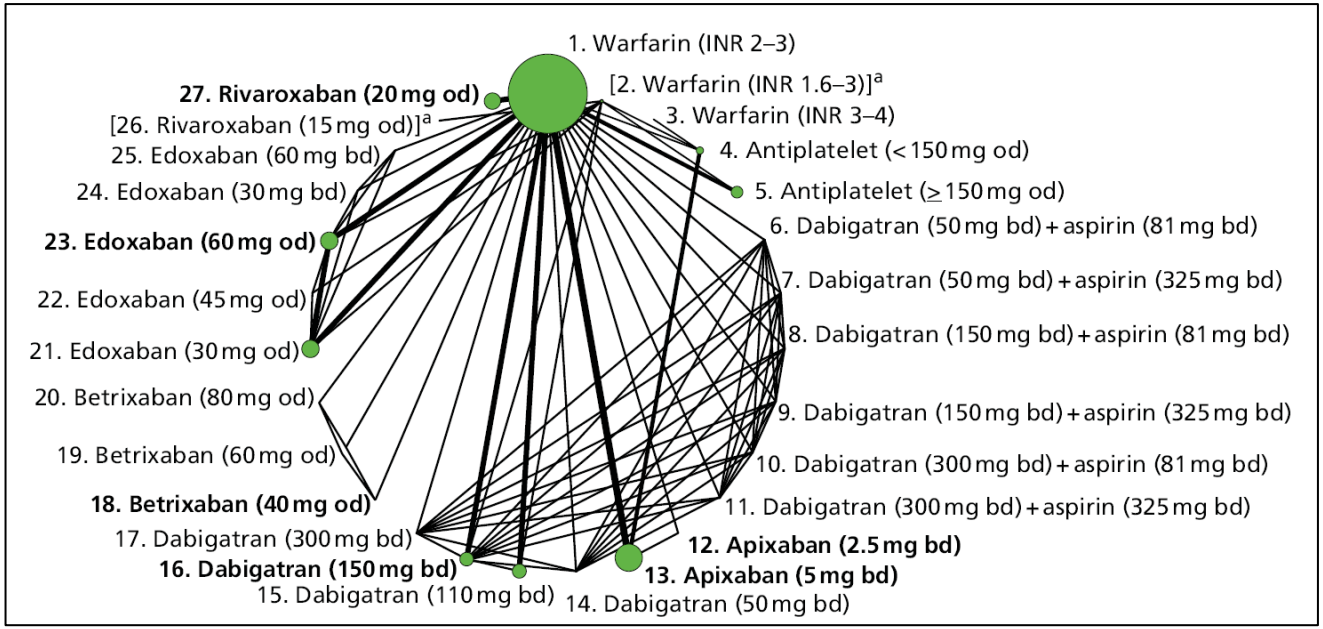


Figure 65: Network plots for stroke or systemic embolism (stroke prevention in AF) [a represents excluded interventions that were included within the sensitivity analyses]

Table 41: Trials included within the NMA and their risk of bias assessment for stroke or SE (stroke prevention in AF)

Study	Interventions	SG	AC	BPP	BOA	IOD	SR
ACTIVE W (265)	1,4	+	+	-	+	+	?
AFASAK (266)	1,4	+	+	-	?	+	?
AFASAK II (267)	1,5	+	?	-	+	+	?
AF-ASA-VKA-CHINA (268)	2,4	?	-	-	-	+	?
AF-DABIG-VKA-JAPAN (269)	2,15,16	?	?	?	?	?	?
AF-EDOX-VKA-ASIA (270)	1,21,23	+	+	-	+	+	+
AF-EDOX-VKA-JAPAN (271)	2,21,22,23	+	?	-	?	+	+
AF-EDOX-VKA-MULTI (272)	1,21,23,24,25	+	+	-	?	+	+
AF-VKA-ASA-CHINA (273)	1,5	+	?	?	?	?	?
ARISTOTLE (274-284)	1,13	+	?	+	+	+	+
ARISTOTLE-J (285)	1,12,13	?	?	-	+	+	+
AVERROES (270, 275, 276, 286)	4,13	+	+	+	+	+	+
BAFTA (286)	1,4	+	+	-	+	+	+
Chinese ATAFS (287)	2,5	?	?	?	?	+	?
ENGAGE AF-TIMI-48 (288, 289)	1,21,23	+	+	+	+	+	+
EXPLORE-Xa (279)	1,18,19,20	?	?	-	+	+	+
J-ROCKET AF (290)	2,26	+	+	+	?	+	+
PATAF (265)	1,5	+	+	?	+	?	+
PETRO (291)	1,6,7,8,9,10,11,14,16,17	?	?	-	?	+	+
RE-LY (292, 293)	1,15,16	+	+	-	+	+	+
ROCKET-AF (294-297)	1,27	+	+	+	?	+	+
SPAF II (298)	1,5	+	?	-	+	+	?
WASPO (299)	1,5	+	+	-	-	+	?

(-) high risk of bias; (+) low risk of bias; (?) unclear risk of bias; (AC) allocation of concealment; (BOA) blinding of outcome; (BPP) blinding of participants and personnel; (IOD) incomplete outcome data; (SG) sequence generation; (SR) selective reporting.

Table 42: Results for stroke or SE (stroke prevention in AF); comparisons with warfarin (INR 2-3)

Comparison with warfarin (INR 2-3)	Direct evidence, OR (95% CI)	NMA, OR (95% CI)
Antiplatelet (< 150mg daily)	1.99 (1.28 to 3.15)	1.88 (1.44 to 2.51)
Antiplatelet (≥ 150mg daily)	1.61 (1.25 to 2.07)	1.61 (1.25 to 2.07)
Apixaban (5mg bd)	0.79 (0.66 to 0.94)	0.79 (0.66 to 0.94)
Dabigatran (110mg bd)	0.90 (0.74 to 1.10)	0.90 (0.74 to 1.10)
Dabigatran (150mg bd)	0.65 (0.52 to 0.81)	0.65 (0.52 to 0.81)
Edoxaban (30mg daily)	1.13 (0.97 to 1.32)	1.13 (0.97 to 1.32)
Edoxaban (60mg daily)	0.86 (0.74 to 1.01)	0.86 (0.74 to 1.01)
Rivaroxaban (20mg daily)	0.88 (0.74 to 1.03)	0.88 (0.74 to 1.03)

Table 43: Results for stroke or SE (stroke prevention in AF); NOACs (licensed doses only)

Licensed NOACs only	Direct evidence, OR (95% CI)	NMA, OR (95% CI)
Dabigatran (150mg bd) vs. apixaban (5mg bd)	-	0.82 (0.62 to 1.08)
Edoxaban (60mg daily) vs. apixaban (5mg bd)	-	1.09 (0.87 to 1.39)
Rivaroxaban (20mg daily) vs. apixaban (5mg bd)	-	1.11 (0.87 to 1.41)
Edoxaban (60mg daily) vs. dabigatran (150mg bd)	-	1.33 (1.02 to 1.75)
Rivaroxaban (20mg daily) vs. dabigatran (150mg bd)	-	1.35 (1.03 to 1.78)
Rivaroxaban (20mg daily) vs. edoxaban (60mg daily)	-	1.01 (0.80 to 1.27)

Table 44: Results for stroke or SE (stroke prevention in AF); comparisons with warfarin (INR 2-3) – sensitivity analysis using HRs instead of ORs

Comparison with warfarin (INR 2-3)	HR (95% CI)
Warfarin (INR 3-4)	0.58 (0.17 to 1.58)
Antiplatelet (< 150mg daily)	1.82 (1.39 to 2.41)
Antiplatelet (≥ 150mg daily)	1.58 (1.23 to 2.02)
Apixaban (5mg bd)	0.79 (0.67 to 0.94)
Dabigatran (110mg bd)	0.91 (0.75 to 1.11)
Dabigatran (150mg bd)	0.66 (0.53 to 0.82)
Edoxaban (30mg daily)	1.13 (0.98 to 1.31)
Edoxaban (60mg daily)	0.87 (0.74 to 1.01)
Rivaroxaban (20mg daily)	0.88 (0.75 to 1.03)

Table 45: Results for stroke or SE (stroke prevention in AF); NOACs (licensed doses only) – sensitivity analysis using HRs instead of ORs

Licensed NOACs only	HR (95% CI)
Dabigatran (150mg bd) vs. apixaban (5mg bd)	0.83 (0.63 to 1.10)
Edoxaban (60mg daily) vs. apixaban (5mg bd)	1.10 (0.87 to 1.38)
Rivaroxaban (20mg daily) vs. apixaban (5mg bd)	1.11 (0.88 to 1.40)
Edoxaban (60mg daily) vs. dabigatran (150mg bd)	1.32 (1.01 to 1.73)
Rivaroxaban (20mg daily) vs. dabigatran (150mg bd)	1.34 (1.02 to 1.76)
Rivaroxaban (20mg daily) vs. edoxaban (60mg daily)	1.01 (0.81 to 1.27)

7.8.4.2 Ischaemic Stroke

Fourteen studies reported on 2,228 ischaemic stroke events, leading to a connected network comparing a total of 15 interventions (see Figure 66). Twelve studies were included in the main analysis, with the remaining two included only in sensitivity analyses. The studies were at mixed risks of bias (see Table 46). There were concerns about lack of blinding of participants for most trials, and about lack of allocation concealment and blinding of outcome assessment in one trial (AF-ASA-VKA-CHINA,(268) only included in sensitivity analyses due to implementation of warfarin within non-standard INR range).

Table 47 shows comparisons of all interventions with warfarin (INR 2–3), suggests that both low- and high-dose antiplatelets increase the risk of ischaemic stroke compared with warfarin (INR 2–3). Among NOACs, there was some evidence that dabigatran (150 mg twice daily) reduces the risk of ischaemic stroke compared with warfarin, whereas edoxaban (30 mg once daily) increases that risk. There was little evidence that the risk of ischaemic stroke differed between licensed doses of NOACs (see Table 48).

In a sensitivity analysis to take into account the differences in duration of follow-up, NMA results were as presented in Table 49 and Table 50. These values are very similar to those for ORs.

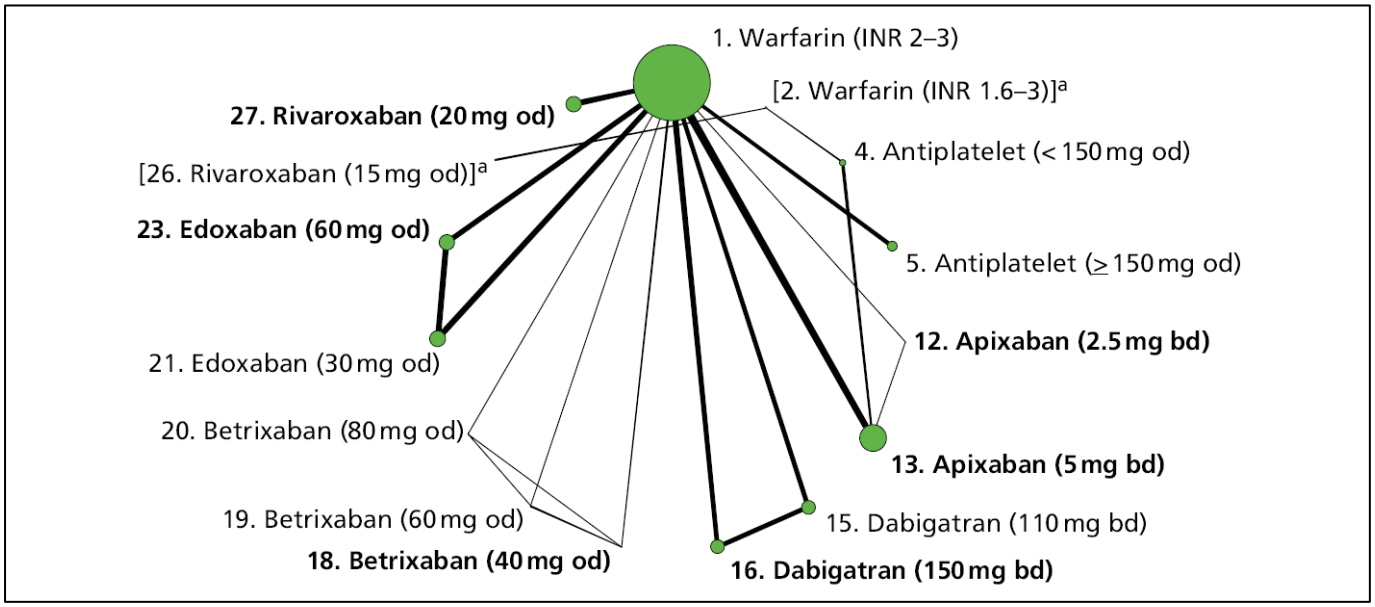


Figure 66: Network plots for ischaemic stroke (stroke prevention in AF) [a represents excluded interventions that were included within the sensitivity analyses]

Table 46: Trials included within the NMA and their risk of bias assessment for ischaemic stroke (stroke prevention in AF)

Study	Interventions	SG	AC	BPP	BOA	IOD	SR
ACTIVE W (265)	1,4	+	+	-	+	+	?
AFASAK II (267)	1,5	+	?	-	+	+	?
AF-ASA-VKA-CHINA (268)	2,4	?	-	-	-	+	?
AF-VKA-ASA-CHINA (273)	1,5	+	?	?	?	?	?
ARISTOTLE (274-284)	1,13	+	?	+	+	+	+
ARISTOTLE-J (285)	1,12,13	?	?	-	+	+	+
AVERROES (270, 275, 276, 286)	4,13	+	+	+	+	+	+
ENGAGE AF-TIMI-48 (288, 289)	1,21,23	+	+	+	+	+	+
EXPLORE-Xa (279)	1,18,19,20	?	?	-	+	+	+
J-ROCKET AF (290)	2,26	+	+	+	?	+	+
PATAF (265)	1,5	+	+	?	+	?	+
RE-LY (292, 293)	1,15,16	+	+	-	+	+	+
ROCKET-AF (294-297)	1,27	+	+	+	?	+	+
SPAF II (298)	1,5	+	?	-	+	+	?

(-) high risk of bias; (+) low risk of bias; (?) unclear risk of bias; (AC) allocation of concealment; (BOA) blinding of outcome; (BPP) blinding of participants and personnel; (IOD) incomplete outcome data; (SG) sequence generation; (SR) selective reporting.

Table 47: Results for ischaemic stroke (stroke prevention in AF); comparisons with warfarin (INR 2-3)

Comparison with warfarin (INR 2-3)	Direct evidence, OR (95% CI)	NMA, OR (95% CI)
Antiplatelet (< 150mg daily)	---	2.52 (1.62 to 3.99)
Antiplatelet (≥ 150mg daily)	2.00 (1.51 to 2.67)	2.00 (1.51 to 2.67)
Apixaban (5mg bd)	0.92 (0.74 to 1.14)	0.92 (0.74 to 1.14)
Dabigatran (110mg bd)	1.14 (0.90 to 1.44)	1.14 (0.90 to 1.44)
Dabigatran (150mg bd)	0.76 (0.58 to 0.98)	0.76 (0.58 to 0.98)
Edoxaban (30mg daily)	1.44 (1.21 to 1.71)	1.44 (1.21 to 1.71)
Edoxaban (60mg daily)	1.01 (0.84 to 1.21)	1.01 (0.84 to 1.21)
Rivaroxaban (20mg daily)	0.93 (0.74 to 1.16)	0.93 (0.74 to 1.16)

Table 48: Results for ischaemic stroke (stroke prevention in AF); NOACs (licensed doses only)

Licensed NOACs only	Direct evidence, OR (95% CI)	NMA, OR (95% CI)
Dabigatran (150mg bd) vs. apixaban (5mg bd)	-	0.83 (0.59 to 1.16)
Edoxaban (60mg daily) vs. apixaban (5mg bd)	-	1.10 (0.83 to 1.46)
Rivaroxaban (20mg daily) vs. apixaban (5mg bd)	-	1.01 (0.74 to 1.38)
Edoxaban (60mg daily) vs. dabigatran (150mg bd)	-	1.33 (0.97 to 1.83)
Rivaroxaban (20mg daily) vs. dabigatran (150mg bd)	-	1.22 (0.87 to 1.73)
Rivaroxaban (20mg daily) vs. edoxaban (60mg daily)	-	0.92 (0.69 to 1.23)

Table 49: Results for ischaemic stroke (stroke prevention in AF); comparisons with warfarin (INR 2-3) – sensitivity analysis using HRs instead of ORs

Comparison with warfarin (INR 2-3)	HR (95% CI)
Antiplatelet (< 150mg daily)	2.46 (1.59 to 3.92)
Antiplatelet (≥ 150mg daily)	1.94 (1.47 to 2.59)
Apixaban (5mg bd)	0.92 (0.75 to 1.15)
Dabigatran (110mg bd)	1.12 (0.89 to 1.42)
Dabigatran (150mg bd)	0.76 (0.59 to 0.99)
Edoxaban (30mg daily)	1.43 (1.22 to 1.69)
Edoxaban (60mg daily)	1.01 (0.84 to 1.20)
Rivaroxaban (20mg daily)	0.92 (0.74 to 1.15)

Table 50: Results for ischaemic stroke (stroke prevention in AF); NOACs (licensed doses only) – sensitivity analysis using HRs instead of ORs

Licensed NOACs only	HR (95% CI)
Dabigatran (150mg bd) vs. apixaban (5mg bd)	0.83 (0.59 to 1.15)
Edoxaban (60mg daily) vs. apixaban (5mg bd)	1.09 (0.83 to 1.44)
Rivaroxaban (20mg daily) vs. apixaban (5mg bd)	1.00 (0.73 to 1.35)
Edoxaban (60mg daily) vs. dabigatran (150mg bd)	1.32 (0.96 to 1.80)
Rivaroxaban (20mg daily) vs. dabigatran (150mg bd)	1.21 (0.86 to 1.70)
Rivaroxaban (20mg daily) vs. edoxaban (60mg daily)	0.92 (0.69 to 1.22)

7.8.4.3 Myocardial Infarction

Fifteen studies reported 1,334 MI events, leading to a network of 16 interventions (see Figure 67). Thirteen studies were included in the main analysis, with the other two included only in sensitivity analyses. The studies were at mixed risks of bias (see Table 51). There were concerns about lack of blinding of participants for most trials, and about lack of allocation concealment and blinding of outcome assessment in some.

Table 52 shows weak evidence that dabigatran (110 mg twice daily), dabigatran (150 mg twice daily) and Edoxaban (30 mg once daily) increase the risk of MI compared with warfarin (INR 2–3), and weak evidence that rivaroxaban (20 mg once daily) decreases risk of MI compared with warfarin (INR 2–3). None of the interventions was superior or inferior to warfarin (INR 2–3). The pairwise comparisons of licensed NOACs, presented in Table 53, show weak evidence that dabigatran (150 mg twice daily) increases the risk of MI compared with Apixaban (5 mg twice daily), and evidence that rivaroxaban (20 mg once daily) reduces the risk of MI compared with Dabigatran (150 mg twice daily). Results were similar in a sensitivity analysis, taking into account the differences in duration of follow-up within and between trials, and the differences in the definition of event used across trials (e.g. total number of events vs. first events only) (see Table 54 and Table 55).

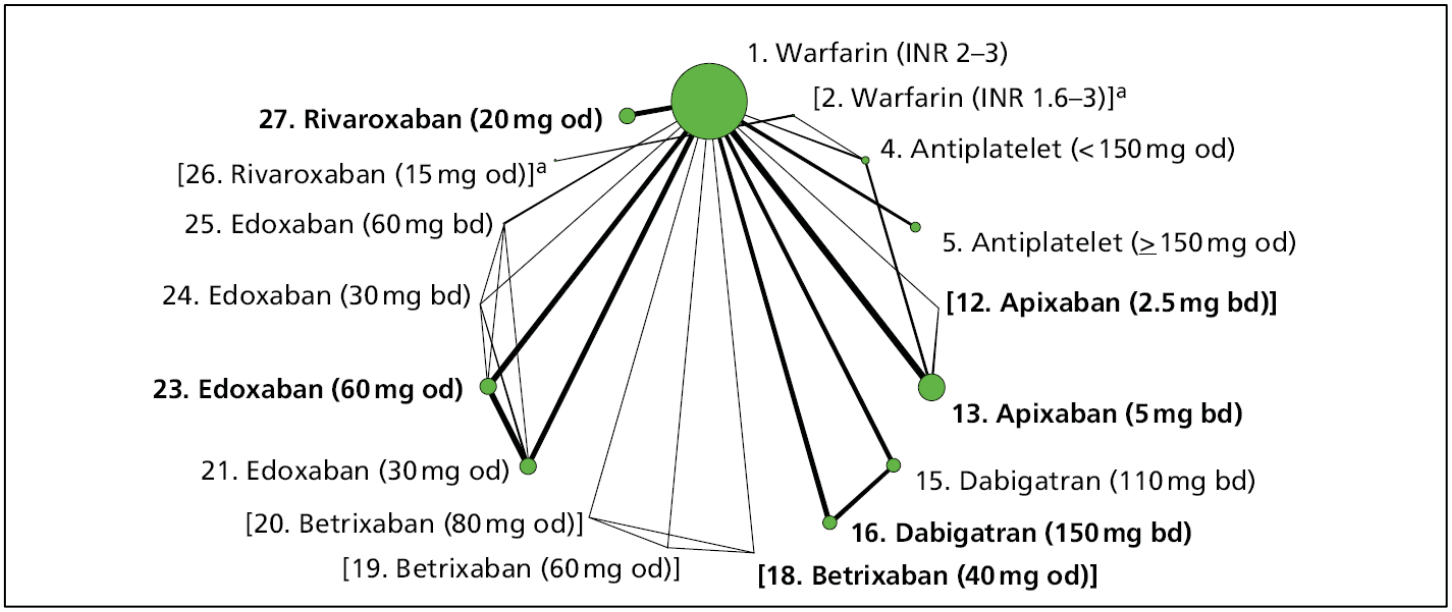


Figure 67: Network plots for MI (stroke prevention in AF) [a represents excluded interventions that were included within the sensitivity analyses]

Table 51: Trials included within the NMA and their risk of bias assessment for MI (stroke prevention in AF)

Study	Interventions	SG	AC	BPP	BOA	IOD	SR
ACTIVE W (265)	1,4	+	+	-	+	+	?
AFASAK II (267)	1,5	+	?	-	+	+	?
AF-ASA-VKA-CHINA (268)	2,4	?	-	-	-	+	?
AF-EDOX-VKA-MULTI (272)	1,21,23,24,25	+	+	-	?	+	+
ARISTOTLE (274-284)	1,13	+	?	+	+	+	+
ARISTOTLE-J (285)	1,12,13	?	?	-	+	+	+
AVERROES (270, 275, 276, 286)	4,13	+	+	+	+	+	+
BAFTA (286)	1,4	+	+	-	+	+	+
ENGAGE AF-TIMI-48 (288, 289)	1,21,23	+	+	+	+	+	+
EXPLORE-Xa (279)	1,18,19,20	?	?	-	+	+	+
J-ROCKET AF (290)	2,26	+	+	+	?	+	+
PATAF (265)	1,5	+	+	?	+	?	+
RE-LY (292, 293)	1,15,16	+	+	-	+	+	+
ROCKET-AF (294-297)	1,27	+	+	+	?	+	+
SPAF II (298)	1,5	+	?	-	-	+	?

(-) high risk of bias; (+) low risk of bias; (?) unclear risk of bias; (AC) allocation of concealment; (BOA) blinding of outcome; (BPP) blinding of participants and personnel; (IOD) incomplete outcome data; (SG) sequence generation; (SR) selective reporting.

Table 52: Results for MI (stroke prevention in AF), comparisons with warfarin (INR 2-3)

Comparison with warfarin (INR 2-3)	Direct evidence, OR (95% CI)	NMA, OR (95% CI)
Antiplatelet (< 150mg daily)	1.00 (0.47 to 2.10)	1.01 (0.64 to 1.61)
Antiplatelet (≥ 150mg daily)	1.38 (0.94 to 2.03)	1.38 (0.94 to 2.03)
Apixaban (5mg bd)	0.87 (0.66 to 1.15)	0.87 (0.66 to 1.15)
Dabigatran (110mg bd)	1.32 (0.97 to 1.79)	1.32 (0.97 to 1.79)
Dabigatran (150mg bd)	1.29 (0.96 to 1.75)	1.29 (0.96 to 1.75)
Edoxaban (30mg daily)	1.22 (0.97 to 1.53)	1.22 (0.97 to 1.53)
Edoxaban (60mg daily)	0.96 (0.75 to 1.22)	0.96 (0.75 to 1.22)
Rivaroxaban (20mg daily)	0.80 (0.61 to 1.04)	0.80 (0.61 to 1.04)

Table 53: Results for MI (stroke prevention in AF), NOACs (licensed doses only)

Licensed NOACs only	Direct evidence, OR (95% CI)	NMA, OR (95% CI)
Dabigatran (150mg bd) vs. apixaban (5mg bd)	-	1.48 (0.98 to 2.22)
Edoxaban (60mg daily) vs. apixaban (5mg bd)	-	1.10 (0.76 to 1.58)
Rivaroxaban (20mg daily) vs. apixaban (5mg bd)	-	0.92 (0.63 to 1.34)
Edoxaban (60mg daily) vs. dabigatran (150mg bd)	-	0.74 (0.50 to 1.09)
Rivaroxaban (20mg daily) vs. dabigatran (150mg bd)	-	0.62 (0.41 to 0.93)
Rivaroxaban (20mg daily) vs. edoxaban (60mg daily)	-	0.84 (0.59 to 1.20)

Table 54: Results for MI (stroke prevention in AF), comparisons with warfarin (INR 2-3) – sensitivity analysis using HRs instead of ORs

Comparison with warfarin (INR 2-3)	HR (95% CI)
Antiplatelet (< 150mg daily)	1.01 (0.64 to 1.61)
Antiplatelet (≥ 150mg daily)	1.36 (0.93 to 2.01)
Apixaban (5mg bd)	0.88 (0.67 to 1.16)
Dabigatran (110mg bd)	1.31 (0.96 to 1.77)
Dabigatran (150mg bd)	1.30 (0.96 to 1.77)
Edoxaban (30mg daily)	1.22 (0.97 to 1.52)
Edoxaban (60mg daily)	0.96 (0.76 to 1.22)
Rivaroxaban (20mg daily)	0.80 (0.62 to 1.04)

Table 55: Results for MI (stroke prevention in AF), NOACs (licensed doses only) – sensitivity analysis using HRs instead of ORs

Licensed NOACs only	HR (95% CI)
Dabigatran (150mg bd) vs. apixaban (5mg bd)	1.48 (0.98 to 2.23)
Edoxaban (60mg daily) vs. apixaban (5mg bd)	1.09 (0.76 to 1.57)
Rivaroxaban (20mg daily) vs. apixaban (5mg bd)	0.91 (0.62 to 1.33)
Edoxaban (60mg daily) vs. dabigatran (150mg bd)	0.74 (0.49 to 1.08)
Rivaroxaban (20mg daily) vs. dabigatran (150mg bd)	0.62 (0.41 to 0.92)
Rivaroxaban (20mg daily) vs. edoxaban (60mg daily)	0.84 (0.59 to 1.19)

7.8.4.4 Major Bleeding

Eighteen studies reporting 4,314 major bleeding events, leading to a network of 24 interventions (see Figure 68). Seventeen studies were included in the main analysis, with the remaining study included only in sensitivity analyses. These studies were at mixed risks of bias (see Table 56). There were concerns about lack of blinding of participants for most trials, and about lack of allocation concealment and blinding of outcome assessment in some.

There was weak evidence that antiplatelet therapy (< 150 mg once daily) reduced major bleeding compared with warfarin (INR 2–3). There was evidence that apixaban (5 mg twice daily), dabigatran (110 mg twice daily), Edoxaban (30 mg once daily) and edoxaban (60 mg once daily) reduced major bleeding risk compared with warfarin (INR 2–3) (see Table 57). Comparisons among licensed doses of NOACs, presented in Table 58, suggest that Dabigatran (150 mg twice daily) increases risk of major bleeding compared with apixaban (5 mg twice daily), whereas rivaroxaban (20 mg once daily) increases risk of major bleeding compared with apixaban (5 mg twice daily) and edoxaban (60 mg once daily).

In a sensitivity analysis to take into account the differences in duration of follow-up, NMA results were as presented in Table 59 and Table 60, and show very similar results. Another sensitivity analysis involved fitting a fixed-effects meta-regression model using the mean TTR for warfarin patients (see Table 35) as a covariate and the mean log-OR from each pairwise comparison (with warfarin as the reference category) as the response variable. No evidence of an effect modification was found according to mean TTR (estimated coefficient 0.04 with 95% CI −0.03 to 0.12 per 1% increase). The model fit indices yielded almost identical values for the models with and without the covariate.

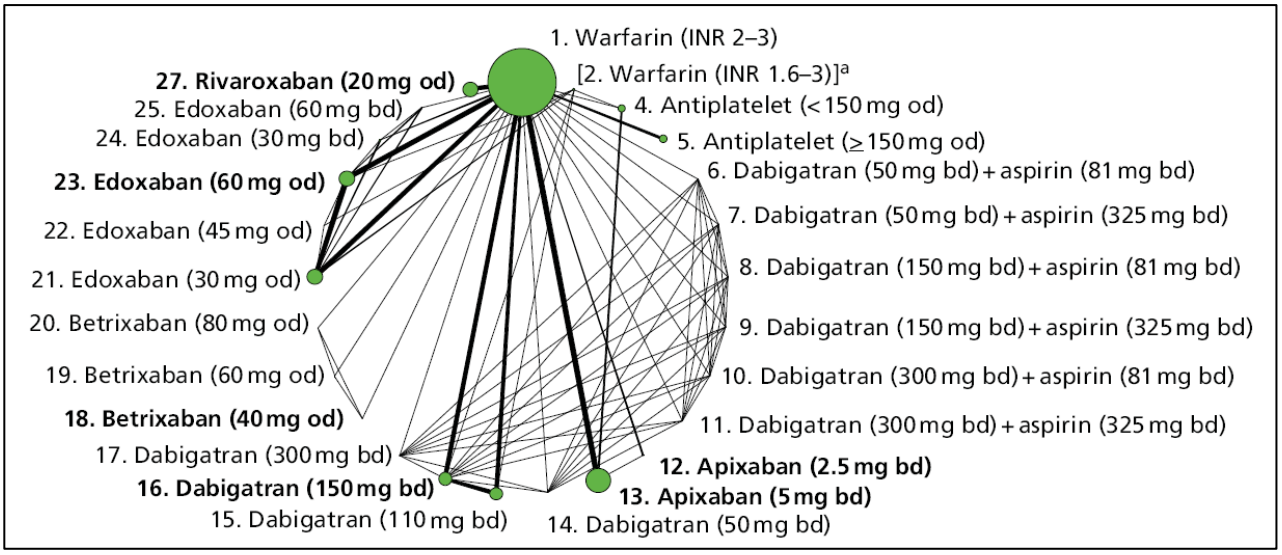


Figure 68: Network plots for major bleeding (stroke prevention in AF) [a represents excluded interventions that were included within the sensitivity analyses]

Table 56: Trials included within the NMA and their risk of bias assessment for major bleeding (stroke prevention in AF)

Study	Interventions	SG	AC	BPP	BOA	IOD	SR
ACTIVE W (265)	1,4	+	+	-	+	+	?
AFASAK II (267)	1,5	+	?	-	+	+	?
AF-ASA-VKA-CHINA (268)	2,4	?	-	-	-	+	?
AF-DABIG-VKA-JAPAN (269)	2,15,16	?	?	?	?	?	?
AF-EDOX-VKA-ASIA (270)	1,21,23	+	+	-	+	+	+
AF-EDOX-VKA-JAPAN (271)	2,21,22,23	+	?	-	?	+	+
AF-EDOX-VKA-MULTI (272)	1,21,23,24,25	+	+	-	+	+	+
AF-VKA-ASA-CHINA (273)	1,5	+	?	-	?	?	?
ARISTOTLE (274-284)	1,13	+	?	+	+	?	+
ARISTOTLE-J (285)	1,12,13	?	?	-	+	+	+
AVERROES (270, 275, 276, 286)	4,13	+	+	+	+	+	+
BAFTA (286)	1,4	+	+	-	+	+	+
ENGAGE AF-TIMI-48 (288, 289)	1,21,23	+	+	+	+	+	+
EXPLORE-Xa (279)	1,18,19,20	?	?	-	+	+	+
PETRO (291)	1,6,7,8,9,10,11,14,16,17	?	?	-	+	+	+
RE-LY (292, 293)	1,15,16	+	+	-	+	+	+
ROCKET-AF (294-297)	1,27	+	+	+	?	+	+
WASPO (299)	1,5	+	+	-	-	+	?

(-) high risk of bias; (+) low risk of bias; (?) unclear risk of bias; (AC) allocation of concealment; (BOA) blinding of outcome; (BPP) blinding of participants and personnel; (IOD) incomplete outcome data; (SG) sequence generation; (SR) selective reporting.

Table 57: Results for major bleeding (stroke prevention in AF), comparisons with warfarin (INR 2-3)

Comparison with warfarin (INR 2-3)	Direct evidence, OR (95% CI)	NMA, OR (95% CI)
Antiplatelet (< 150mg daily)	1.00 (0.56 to 1.77)	0.75 (0.52 to 1.06)
Antiplatelet (≥ 150mg daily)	1.07 (0.82 to 1.42)	1.07 (0.82 to 1.42)
Apixaban (5mg bd)	0.71 (0.61 to 0.81)	0.71 (0.61 to 0.81)
Dabigatran (110mg bd)	0.80 (0.69 to 0.93)	0.80 (0.69 to 0.93)
Dabigatran (150mg bd)	0.94 (0.81 to 1.08)	0.94 (0.81 to 1.08)
Edoxaban (30mg daily)	0.46 (0.40 to 0.54)	0.46 (0.40 to 0.54)
Edoxaban (60mg daily)	0.78 (0.69 to 0.90)	0.78 (0.69 to 0.90)
Rivaroxaban (20mg daily)	1.03 (0.89 to 1.18)	1.03 (0.89 to 1.18)

Table 58: Results for major bleeding (stroke prevention in AF), NOACs (licensed doses only)

Licensed NOACs only	Direct evidence, OR (95% CI)	NMA, OR (95% CI)
Dabigatran (150mg bd) vs. apixaban (5mg bd)	-	1.33 (1.09 to 1.62)
Edoxaban (60mg daily) vs. apixaban (5mg bd)	-	1.11 (0.92 to 1.35)
Rivaroxaban (20mg daily) vs. apixaban (5mg bd)	-	1.45 (1.19 to 1.78)
Edoxaban (60mg daily) vs. dabigatran (150mg bd)	-	0.84 (0.69 to 1.02)
Rivaroxaban (20mg daily) vs. dabigatran (150mg bd)	-	1.10 (0.90 to 1.34)
Rivaroxaban (20mg daily) vs. edoxaban (60mg daily)	-	1.31 (1.07 to 1.59)

Table 59: Results for major bleeding (stroke prevention in AF), comparisons with warfarin (INR 2-3) – sensitivity analysis using HRs instead of ORs

Comparison with warfarin (INR 2-3)	HR (95% CI)
Antiplatelet (< 150mg daily)	0.76 (0.53 to 1.08)
Antiplatelet (≥ 150mg daily)	1.07 (0.82 to 1.41)
Apixaban (5mg bd)	0.72 (0.62 to 0.82)
Dabigatran (110mg bd)	0.81 (0.70 to 0.93)
Dabigatran (150mg bd)	0.94 (0.82 to 1.07)
Edoxaban (30mg daily)	0.47 (0.41 to 0.55)
Edoxaban (60mg daily)	0.79 (0.70 to 0.90)
Rivaroxaban (20mg daily)	1.02 (0.89 to 1.18)

Table 60: Results for major bleeding (stroke prevention in AF), NOACs (licensed doses only) – sensitivity analysis using HRs instead of ORs

Licensed NOACs only	HR (95% CI)
Dabigatran (150mg bd) vs. apixaban (5mg bd)	1.31 (1.08 to 1.59)
Edoxaban (60mg daily) vs. apixaban (5mg bd)	1.10 (0.91 to 1.33)
Rivaroxaban (20mg daily) vs. apixaban (5mg bd)	1.43 (1.17 to 1.75)
Edoxaban (60mg daily) vs. dabigatran (150mg bd)	0.84 (0.70 to 1.02)
Rivaroxaban (20mg daily) vs. dabigatran (150mg bd)	1.09 (0.90 to 1.33)
Rivaroxaban (20mg daily) vs. edoxaban (60mg daily)	1.30 (1.07 to 1.57)

7.8.4.5 Clinically Relevant Bleeding

Twelve studies reporting 9,556 clinically relevant bleeding (CRB) events, leading to a network of 23 interventions (see Figure 69). Eleven studies were included in the main analysis, with the remaining study included only in sensitivity analyses. These studies were at mixed risks of bias (see Table 61), the concerns being due to a lack of blinding of participants for most trials.

Results presented in Table 62 suggest that antiplatelet therapy (< 150 mg once daily) reduces CRB compared with warfarin (INR 2–3). Note that the licensed dose for antiplatelet therapy for AF is ≥ 150 mg once daily: no studies provided data for that dose for CRB. Among NOACs, there was evidence that apixaban (5 mg twice daily), edoxaban (30 mg once daily) and edoxaban (60 mg once daily) reduce CRB compared with warfarin (INR 2–3). However, edoxaban (30 mg twice daily) and edoxaban (60 mg twice daily) increased CRB compared with warfarin (INR 2–3). Among licensed NOACs (see Table 63), there was evidence that edoxaban (60 mg once daily) and rivaroxaban (20 mg once daily) increase CRB compared with apixaban (5 mg twice daily) and that rivaroxaban (20 mg once daily) increased CRB compared with edoxaban (60 mg once daily).

Results from a supplementary analysis of HRs rather than ORs show very similar results (see Table 64 and Table 65).

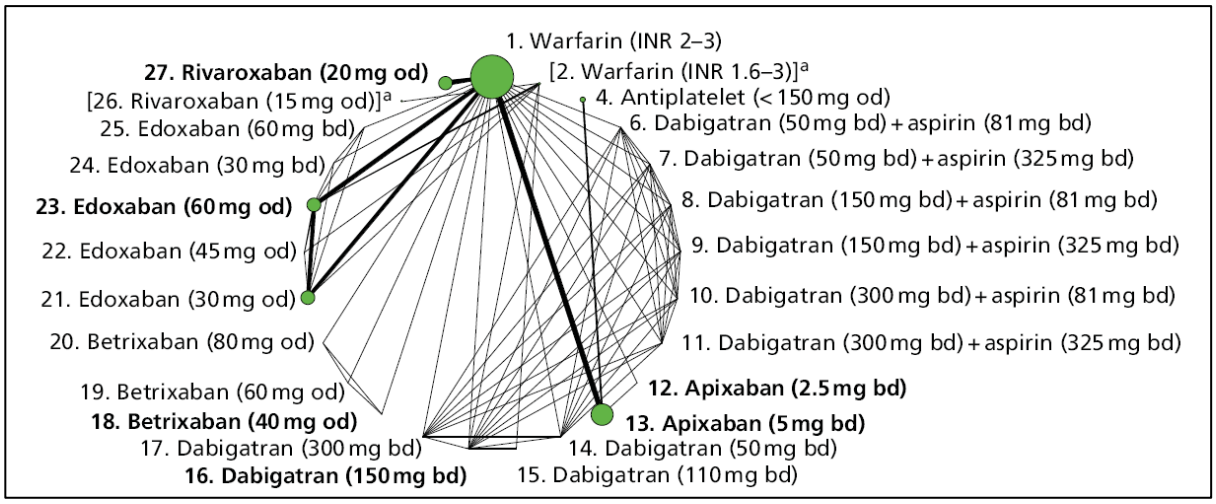


Figure 69: Network plots for CRB (stroke prevention in AF) [a represents excluded interventions that were included within the sensitivity analyses]

Table 61: Trials included within the NMA and their risk of bias assessment for CRB (stroke prevention in AF)

Study	Interventions	SG	AC	BPP	BOA	IOD	SR
AF-DABIG-VKA-JAPAN (269)	2,15,16	?	?	?	?	?	?
AF-EDOX-VKA-ASIA (270)	1,21,23	+	+	-	+	+	+
AF-EDOX-VKA-JAPAN (271)	2,21,22,23	+	?	-	?	+	+
AF-EDOX-VKA-MULTI (272)	1,21,23,24,25	+	+	-	+	+	+
ARISTOTLE (274-284)	1,13	+	?	+	+	?	+
ARISTOTLE-J (285)	1,12,13	?	?	-	+	+	+
AVERROES (270, 275, 276, 286)	4,13	+	+	+	+	+	+
ENGAGE AF-TIMI-48 (288, 289)	1,21,23	+	+	+	+	+	+
EXPLORE-Xa (279)	1,18,19,20	?	?	-	+	+	+
J-ROCKET AF (290)	2,26	+	+	+	?	+	+
PETRO (291)	1,6,7,8,9,10,11,14,16,17	?	?	-	+	+	+
ROCKET-AF (294-297)	1,27	+	+	+	?	+	+

(-) high risk of bias; (+) low risk of bias; (?) unclear risk of bias; (AC) allocation of concealment; (BOA) blinding of outcome; (BPP) blinding of participants and personnel; (IOD) incomplete outcome data; (SG) sequence generation; (SR) selective reporting.

Table 62: Results for CRB (stroke prevention in AF), comparisons with warfarin (INR 2-3)

Comparison with warfarin (INR 2-3)	Direct evidence, OR (95% CI)	NMA, OR (95% CI)
Antiplatelet (< 150mg daily)	---	0.59 (0.45 to 0.77)
Antiplatelet (≥ 150mg daily)	0.67 (0.60 to 0.75)	0.67 (0.60 to 0.75)
Apixaban (5mg bd)	0.59 (0.54 to 0.64)	0.59 (0.54 to 0.64)
Dabigatran (110mg bd)	1.09 (0.37 to 3.04)	1.09 (0.37 to 3.04)
Dabigatran (150mg bd)	0.84 (0.77 to 0.90)	0.84 (0.77 to 0.90)
Edoxaban (30mg daily)	1.97 (1.04 to 3.67)	1.97 (1.04 to 3.67)
Edoxaban (60mg daily)	2.76 (1.46 to 5.17)	2.76 (1.46 to 5.17)
Rivaroxaban (20mg daily)	1.03 (0.95 to 1.11)	1.03 (0.95 to 1.11)

Table 63: Results for CRB (stroke prevention in AF), NOACs (licensed doses only)

Licensed NOACs only	Direct evidence, OR (95% CI)	NMA, OR (95% CI)
Edoxaban (60mg daily) vs. apixaban (5mg bd)	-	1.24 (1.09 to 1.42)
Rivaroxaban (20mg daily) vs. apixaban (5mg bd)	-	1.53 (1.33 to 1.75)
Rivaroxaban (20mg daily) vs. edoxaban (60mg daily)	-	1.23 (1.10 to 1.37)

Table 64: Results for CRB (stroke prevention in AF), comparisons with warfarin (INR 2-3) – sensitivity analysis using HRs instead of ORs

Comparison with warfarin (INR 2-3)	HR (95% CI)
Antiplatelet (< 150mg daily)	0.59 (0.46 to 0.76)
Apixaban (5mg bd)	0.67 (0.60 to 0.75)
Edoxaban (30mg daily)	0.59 (0.55 to 0.64)
Edoxaban (45mg daily)	1.09 (0.37 to 3.01)
Edoxaban (60mg daily)	0.83 (0.77 to 0.90)
Edoxaban (30mg bd)	1.98 (1.05 to 3.71)
Edoxaban (60mg bd)	2.78 (1.46 to 5.20)
Rivaroxaban (20mg daily)	1.03 (0.95 to 1.11)

Table 65: Results for CRB (stroke prevention in AF), NOACs (licensed doses only) – sensitivity analysis using HRs instead of ORs

Licensed NOACs only	HR (95% CI)
Edoxaban (60mg daily) vs. apixaban (5mg bd)	1.24 (1.09 to 1.42)
Rivaroxaban (20mg daily) vs. apixaban (5mg bd)	1.53 (1.33 to 1.74)
Rivaroxaban (20mg daily) vs. edoxaban (60mg daily)	1.23 (1.10 to 1.37)

7.8.4.6 Intracranial Bleeding

Eight studies reporting a total of 757 intracranial bleeds, leading to a network of 10 interventions (see Figure 70). Seven studies were included in the main analysis, with the remaining study included only in sensitivity analyses. These studies were at mixed risks of bias (see Table 66), the concerns being due to a lack of blinding of participants and, in one study, lack of blinding of outcome assessment.

There was strong evidence that Apixaban (5mg twice daily), dabigatran (110mg twice daily), dabigatran (150mg twice daily), edoxaban (30mg once daily), edoxaban (60 mg once daily) and rivaroxaban (20 mg once daily) reduced risk of intracranial bleeding compared with warfarin (INR 2–3) (see Table 67). For each of these NOAC doses except rivaroxaban (20 mg once daily), the estimated reduction in risk was > 50%. There was weak evidence that risk of intracranial bleeding was increased for rivaroxaban (20 mg once daily) compared with apixaban (5 mg twice daily), dabigatran (150 mg twice daily) and edoxaban (60 mg once daily) (see Table 68). Analysing HRs rather than ORs led to similar results (see Table 69 and Table 70).

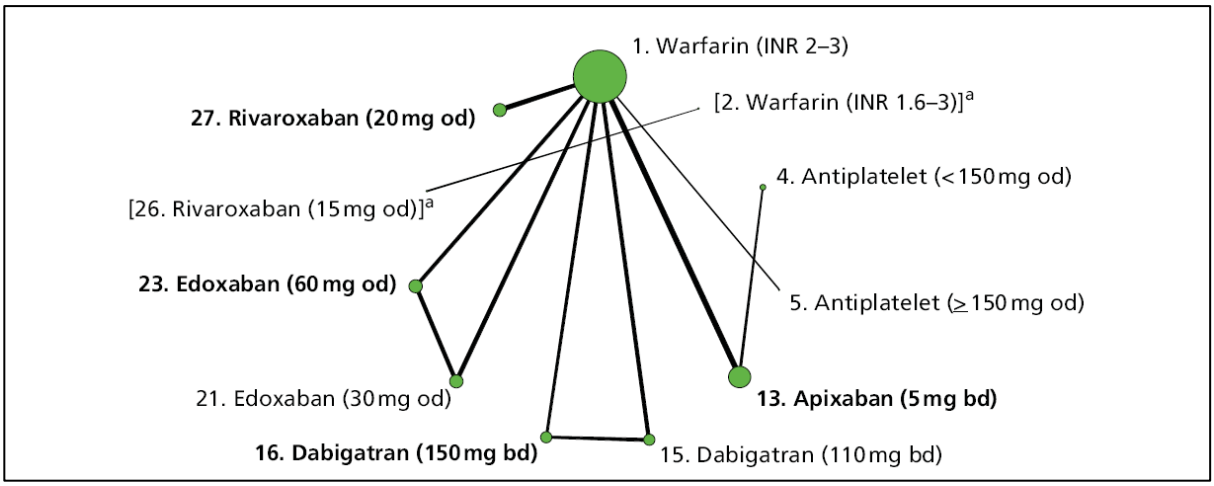


Figure 70: Network plots for intracranial bleeding (stroke prevention in AF) [a represents excluded interventions that were included within the sensitivity analyses]

Table 66: Trials included within the NMA and their risk of bias assessment for intracranial bleeding (stroke prevention in AF)

Study	Interventions	SG	AC	BPP	BOA	IOD	SR
AFASAK II (267)	1,5	+	?	-	+	+	?
ARISTOTLE (274-284)	1,13	+	?	+	+	?	+
AVERROES (270, 275, 276, 286)	4,13	+	+	+	+	+	+
ENGAGE AF-TIMI-48 (288, 289)	1,21,23	+	+	+	+	+	+
RE-LY (292, 293)	1,15,16	+	+	-	+	+	+
J-ROCKET AF (290)	2,26	+	+	+	?	+	+
ROCKET-AF (294-297)	1,27	+	+	+	?	+	+
SPAF II (298)	1,5	+	?	-	-	+	?

(-) high risk of bias; (+) low risk of bias; (?) unclear risk of bias; (AC) allocation of concealment; (BOA) blinding of outcome; (BPP) blinding of participants and personnel; (IOD) incomplete outcome data; (SG) sequence generation; (SR) selective reporting.

Table 67: Results for intracranial bleeding (stroke prevention in AF), comparisons with warfarin (INR 2-3)

Comparison with warfarin (INR 2-3)	Direct evidence, OR (95% CI)	NMA, OR (95% CI)
Antiplatelet (< 150mg daily)	---	0.50 (0.21 to 1.23)
Antiplatelet (≥ 150mg daily)	0.39 (0.13 to 0.98)	0.39 (0.13 to 0.98)
Apixaban (5mg bd)	0.42 (0.30 to 0.58)	0.42 (0.30 to 0.58)
Dabigatran (110mg bd)	0.31 (0.19 to 0.47)	0.31 (0.19 to 0.47)
Dabigatran (150mg bd)	0.40 (0.27 to 0.59)	0.40 (0.27 to 0.59)
Edoxaban (30mg daily)	0.31 (0.21 to 0.43)	0.31 (0.21 to 0.43)
Edoxaban (60mg daily)	0.46 (0.33 to 0.62)	0.46 (0.33 to 0.62)
Rivaroxaban (20mg daily)	0.65 (0.46 to 0.91)	0.65 (0.46 to 0.91)

Table 68: Results for intracranial bleeding (stroke prevention in AF), NOACs (licensed doses only)

Licensed NOACs only	Direct evidence, OR (95% CI)	NMA, OR (95% CI)
Dabigatran (150mg bd) vs. apixaban (5mg bd)	-	0.96 (0.58 to 1.60)
Edoxaban (60mg daily) vs. apixaban (5mg bd)	-	1.09 (0.69 to 1.70)
Rivaroxaban (20mg daily) vs. apixaban (5mg bd)	-	1.55 (0.97 to 2.49)
Edoxaban (60mg daily) vs. dabigatran (150mg bd)	-	1.13 (0.69 to 1.87)
Rivaroxaban (20mg daily) vs. dabigatran (150mg bd)	-	1.61 (0.96 to 2.72)
Rivaroxaban (20mg daily) vs. edoxaban (60mg daily)	-	1.43 (0.90 to 2.26)

Table 69: Results for intracranial bleeding (stroke prevention in AF), comparisons with warfarin (INR 2-3) – sensitivity analysis using HRs instead of ORs

Comparison with warfarin (INR 2-3)	HR (95% CI)
Antiplatelet (< 150mg daily)	0.50 (0.21 to 1.20)
Antiplatelet (≥ 150mg daily)	0.39 (0.14 to 0.97)
Apixaban (5mg bd)	0.42 (0.30 to 0.58)
Dabigatran (110mg bd)	0.31 (0.19 to 0.46)
Dabigatran (150mg bd)	0.41 (0.27 to 0.59)
Edoxaban (30mg daily)	0.31 (0.21 to 0.43)
Edoxaban (60mg daily)	0.46 (0.34 to 0.62)
Rivaroxaban (20mg daily)	0.66 (0.47 to 0.91)

Table 70: Results for intracranial bleeding (stroke prevention in AF), NOACs (licensed doses only) – sensitivity analysis using HRs instead of ORs

Licensed NOACs only	HR (95% CI)
Dabigatran (150mg bd) vs. apixaban (5mg bd)	0.97 (0.57 to 1.58)
Edoxaban (60mg daily) vs. apixaban (5mg bd)	1.09 (0.70 to 1.71)
Rivaroxaban (20mg daily) vs. apixaban (5mg bd)	1.55 (0.97 to 2.48)
Edoxaban (60mg daily) vs. dabigatran (150mg bd)	1.13 (0.70 to 1.87)
Rivaroxaban (20mg daily) vs. dabigatran (150mg bd)	1.62 (0.96 to 2.74)
Rivaroxaban (20mg daily) vs. edoxaban (60mg daily)	1.43 (0.91 to 2.25)

7.8.4.7 All-Cause Mortality

Eighteen studies reported 6,479 all-cause mortality events, leading to a network of 15 interventions (see Figure 71). Fifteen studies were included in the primary analysis, with the remaining study included only in sensitivity analyses. These studies were at mixed risks of bias (see Table 71), the concerns being due to a lack of blinding of participants for most studies, and about lack of allocation concealment and blinding of outcome assessment in some studies.

Table 72 suggests that all NOAC doses with comparisons that were not imprecisely estimated [apixaban (5 mg twice daily), dabigatran (110 mg twice daily), dabigatran (150 mg twice daily), edoxaban (30 mg once daily), edoxaban (60 mg once daily) and rivaroxaban (20 mg once daily)] were associated with a reduced risk of all-cause mortality compared with warfarin (INR 2–3). There was little evidence that the risk of all-cause mortality differed between licensed doses of NOACs (see Table 73). Analysing HRs rather than ORs produced similar results (see Table 74 and Table 75).

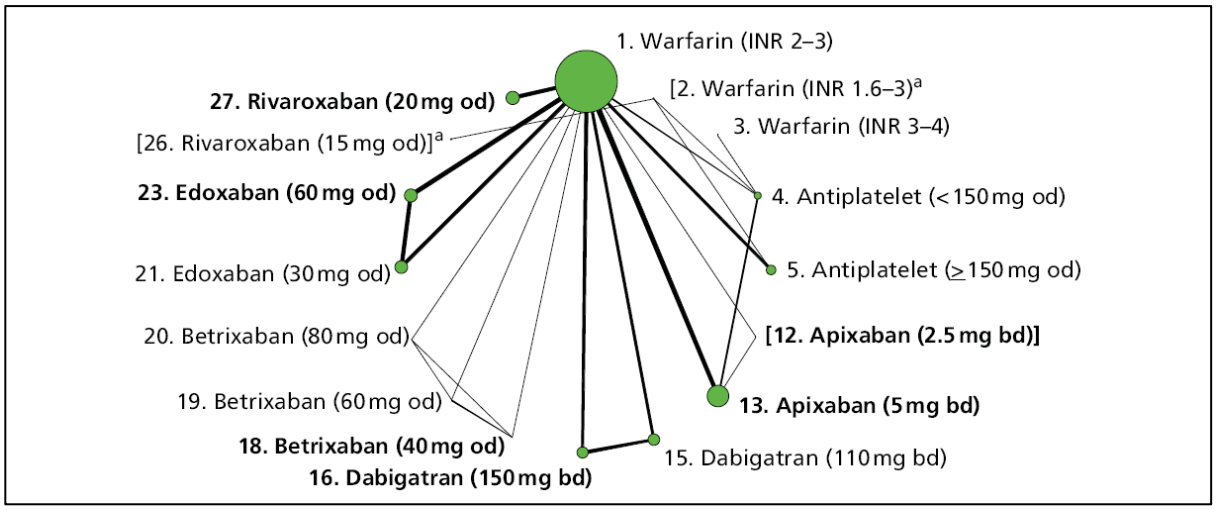


Figure 71: Network plots for all-cause mortality (stroke prevention in AF) [a represents excluded interventions that were included within the sensitivity analyses]

Table 71: Trials included within the NMA and their risk of bias assessment for all-cause mortality (stroke prevention in AF)

Study	Interventions	SG	AC	BPP	BOA	IOD	SR
ACTIVE W (265)	1,4	+	+	-	+	+	?
AFASAK (266)	1,4	+	+	-	?	+	?
AFASAK II (267)	1,5	+	?	-	+	+	?
AF-ASA-VKA-CHINA (268)	2,4	?	-	-	-	+	?
AF-VKA-ASA-CHINA (273)	1,5	+	?	-	?	?	?
ARISTOTLE (274-284)	1,13	+	?	+	+	+	+
ARISTOTLE-J (285)	1,12,13	?	?	-	+	+	+
AVERROES (270, 275, 276, 286)	4,13	+	+	+	+	+	+
BAFTA (286)	1,4	+	+	-	+	+	+
Chinese ATAFS (287)	2,5	?	?	?	?	+	?
ENGAGE AF-TIMI-48 (288, 289)	1,21,23	+	+	+	+	+	+
EXPLORE-Xa (279)	1,18,19,20	?	?	-	+	+	+
J-ROCKET AF (290)	2,26	+	+	+	?	+	+
PATAF (265)	1,5	+	+	?	+	?	+
RE-LY (292, 293)	1,15,16	+	+	-	+	+	+
ROCKET-AF (294-297)	1,27	+	+	+	+	+	+
SPAF II (298)	1,5	+	?	-	-	+	?
WASPO (299)	1,5	+	+	-	-	+	?

(-) high risk of bias; (+) low risk of bias; (?) unclear risk of bias; (AC) allocation of concealment; (BOA) blinding of outcome; (BPP) blinding of participants and personnel; (IOD) incomplete outcome data; (SG) sequence generation; (SR) selective reporting.

Table 72: Results for all-cause mortality (stroke prevention in AF), comparisons with warfarin (INR 2-3)

Comparison with warfarin (INR 2-3)	Direct evidence, OR (95% CI)	NMA, OR (95% CI)
Antiplatelet (< 150mg daily)	1.02 (0.75 to 1.38)	1.08 (0.88 to 1.33)
Antiplatelet (≥ 150mg daily)	1.04 (0.87 to 1.25)	1.04 (0.87 to 1.25)
Apixaban (5mg bd)	0.88 (0.79 to 0.98)	0.88 (0.79 to 0.98)
Dabigatran (110mg bd)	0.91 (0.80 to 1.04)	0.91 (0.80 to 1.04)
Dabigatran (150mg bd)	0.88 (0.77 to 1.01)	0.88 (0.77 to 1.01)
Edoxaban (30mg daily)	0.86 (0.78 to 0.96)	0.86 (0.78 to 0.96)
Edoxaban (60mg daily)	0.91 (0.82 to 1.01)	0.91 (0.82 to 1.01)
Rivaroxaban (20mg daily)	0.83 (0.69 to 1.00)	0.83 (0.69 to 1.00)

Table 73: Results for all-cause mortality (stroke prevention in AF), NOACs (licensed doses only)

Licensed NOACs only	Direct evidence, OR (95% CI)	NMA, OR (95% CI)
Dabigatran (150mg bd) vs. apixaban (5mg bd)	-	1.00 (0.84 to 1.19)
Edoxaban (60mg daily) vs. apixaban (5mg bd)	-	1.03 (0.89 to 1.20)
Rivaroxaban (20mg daily) vs. apixaban (5mg bd)	-	0.94 (0.76 to 1.17)
Edoxaban (60mg daily) vs. dabigatran (150mg bd)	-	1.03 (0.87 to 1.22)
Rivaroxaban (20mg daily) vs. dabigatran (150mg bd)	-	0.94 (0.74 to 1.18)
Rivaroxaban (20mg daily) vs. edoxaban (60mg daily)	-	0.91 (0.73 to 1.13)

Table 74: Results for all-cause mortality (stroke prevention in AF), comparisons with warfarin (INR 2-3) – sensitivity analysis using HRs instead of ORs

Comparison with warfarin (INR 2-3)	HR (95% CI)
Antiplatelet (< 150mg daily)	1.07 (0.88 to 1.30)
Antiplatelet (≥ 150mg daily)	1.04 (0.87 to 1.24)
Apixaban (5mg bd)	0.89 (0.80 to 0.99)
Dabigatran (110mg bd)	0.91 (0.80 to 0.99)
Dabigatran (150mg bd)	0.89 (0.78 to 1.01)
Edoxaban (30mg daily)	0.88 (0.80 to 0.97)
Edoxaban (60mg daily)	0.92 (0.83 to 1.02)
Rivaroxaban (20mg daily)	0.83 (0.69 to 1.00)

Table 75: Results for all-cause mortality (stroke prevention in AF), NOACs (licensed doses only) – sensitivity analysis using HRs instead of ORs

Licensed NOACs only	HR (95% CI)
Dabigatran (150mg bd) vs. apixaban (5mg bd)	1.00 (0.85 to 1.18)
Edoxaban (60mg daily) vs. apixaban (5mg bd)	1.03 (0.90 to 1.20)
Rivaroxaban (20mg daily) vs. apixaban (5mg bd)	0.94 (0.76 to 1.15)
Edoxaban (60mg daily) vs. dabigatran (150mg bd)	1.03 (0.88 to 1.22)
Rivaroxaban (20mg daily) vs. dabigatran (150mg bd)	0.93 (0.75 to 1.17)
Rivaroxaban (20mg daily) vs. edoxaban (60mg daily)	0.90 (0.73 to 1.11)

7.8.5 Summary of results

Results from the NMA suggest that a number of licensed doses of NOACs reduce the risk of the outcome of stroke or SE, major bleeding, CRB, intracranial bleeding and all-cause mortality compared with the reference treatment, warfarin (doses adjusted to maintain INR 2-3). There was evidence that edoxaban increased CRB compared with warfarin (INR 2-3). Risk of MI appeared higher for some NOACs than for warfarin (INR 2–3). Comparisons for some licensed NOAC doses, such as apixaban (2.5 mg twice daily) and betrixaban (40 mg once daily), could not be estimated precisely.

Due to local protocols, several studies conducted in Asian countries considered a lower INR range for warfarin interventions in elderly patients. These were excluded from the main analysis but were included (merged with the reference treatment, warfarin INR 2–3) as a second sensitivity analysis for each outcome. This permitted the incorporation of a non-licensed dose of rivaroxaban (15 mg once daily) that was included in the J-ROCKET AF trial showing a reduced risk of stroke compared with warfarin (INR 1.6–3), with a median OR of 0.49 (95% CI 0.24 to 0.99). Apart from this, results showed the same trends as described above.

The dose range for the antiplatelet arm in the AVERROES trial was unusually wide (81–324 mg once daily). Because some of the patients had received a dose that was below UK standard, it was agreed between the clinical collaborators of this project to merge this with the antiplatelets (< 150 mg once daily) node for the primary analysis. In a further sensitivity analysis for each outcome, this trial was excluded. Again, the results were not substantially different from those presented above.

With regard to appraisal of the NMA model, no instances of lack of convergence among the Markov chains, poor model fit or poor model inconsistency were identified.

7.8.5.1 Trade-off Analysis

The trade-off analysis, presented in the form of rank-o-grams, plotted the probability that each of the licensed interventions for AF is ranked best, second best, and so on, for preventing each outcome, are displayed in Figure 72 and Figure 73, for efficacy and safety, respectively.

The non-NOAC interventions (warfarin, INR 2–3) and antiplatelet therapy (aspirin/clopidogrel, ≥ 150 mg once daily) were ranked least effective for stroke or SE and ischaemic stroke and were not among the best three interventions for any of the outcomes. Warfarin (INR 2–3) was also ranked as the least effective intervention to reduce the risk of intracranial bleeding. Among the licensed NOACs, apixaban (5 mg twice daily) was ranked as among the best interventions for major bleeding, intracranial bleeding, all-cause mortality, stroke or SE, ischaemic stroke and MI. Edoxaban (60 mg once daily) was ranked second for major bleeding and all-cause mortality. Except for all-cause mortality and MI, outcomes for rivaroxaban (20 mg once daily) were ranked less highly than those for apixaban (5 mg twice daily), dabigatran (150 mg twice daily) and edoxaban (60 mg once daily).

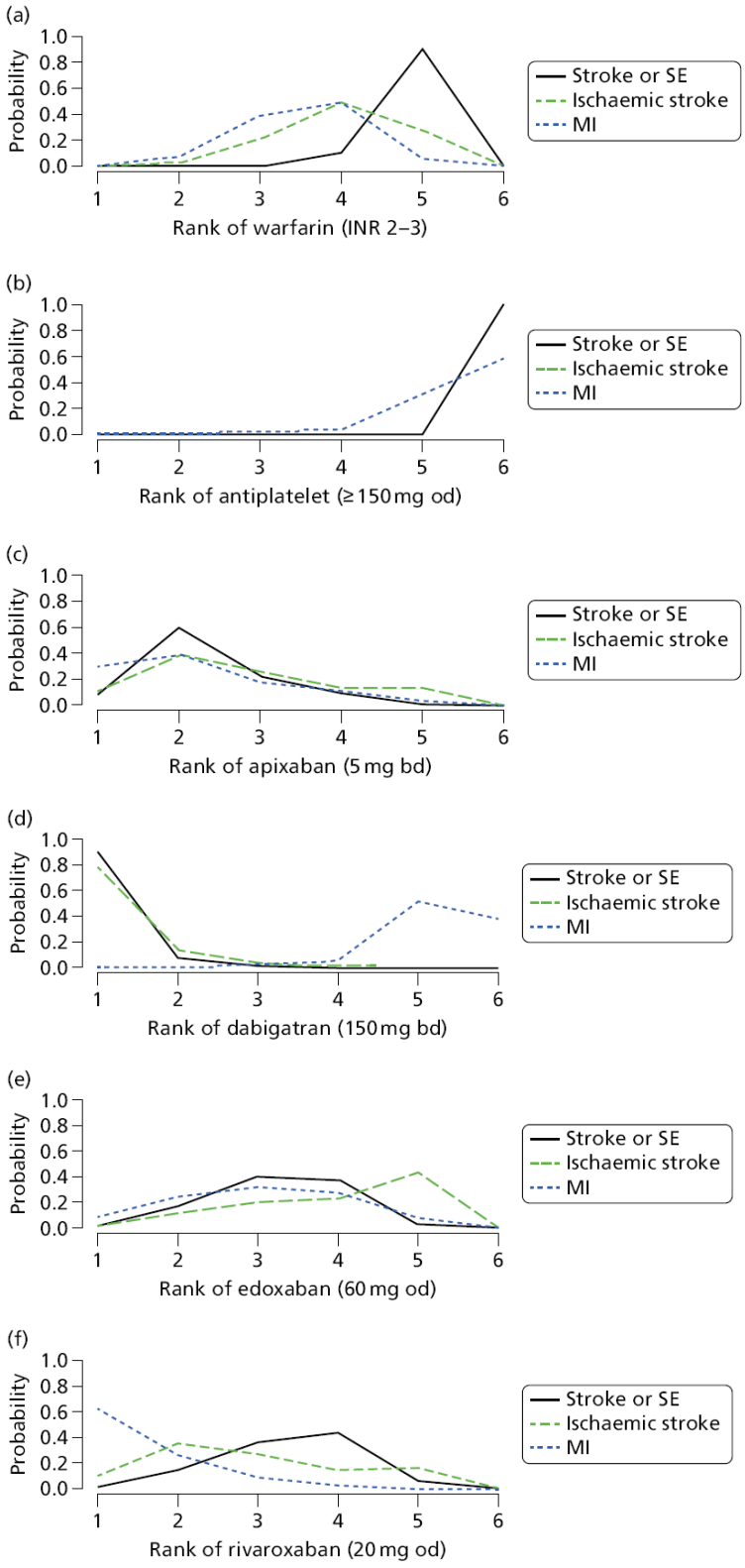


Figure 72: Ranking of efficacy (Rank-o-gram) for licensed interventions examined in stroke prevention in AF

Ranking indicates the probability to be the best treatment, the second best, the third best, and so on, among the NOACs (and warfarin). A higher probability (y-axis) at a rank of 1(x-axis) indicated the most favourable treatment for the plotted endpoint (stroke or SE, ischaemic stroke, or MI).

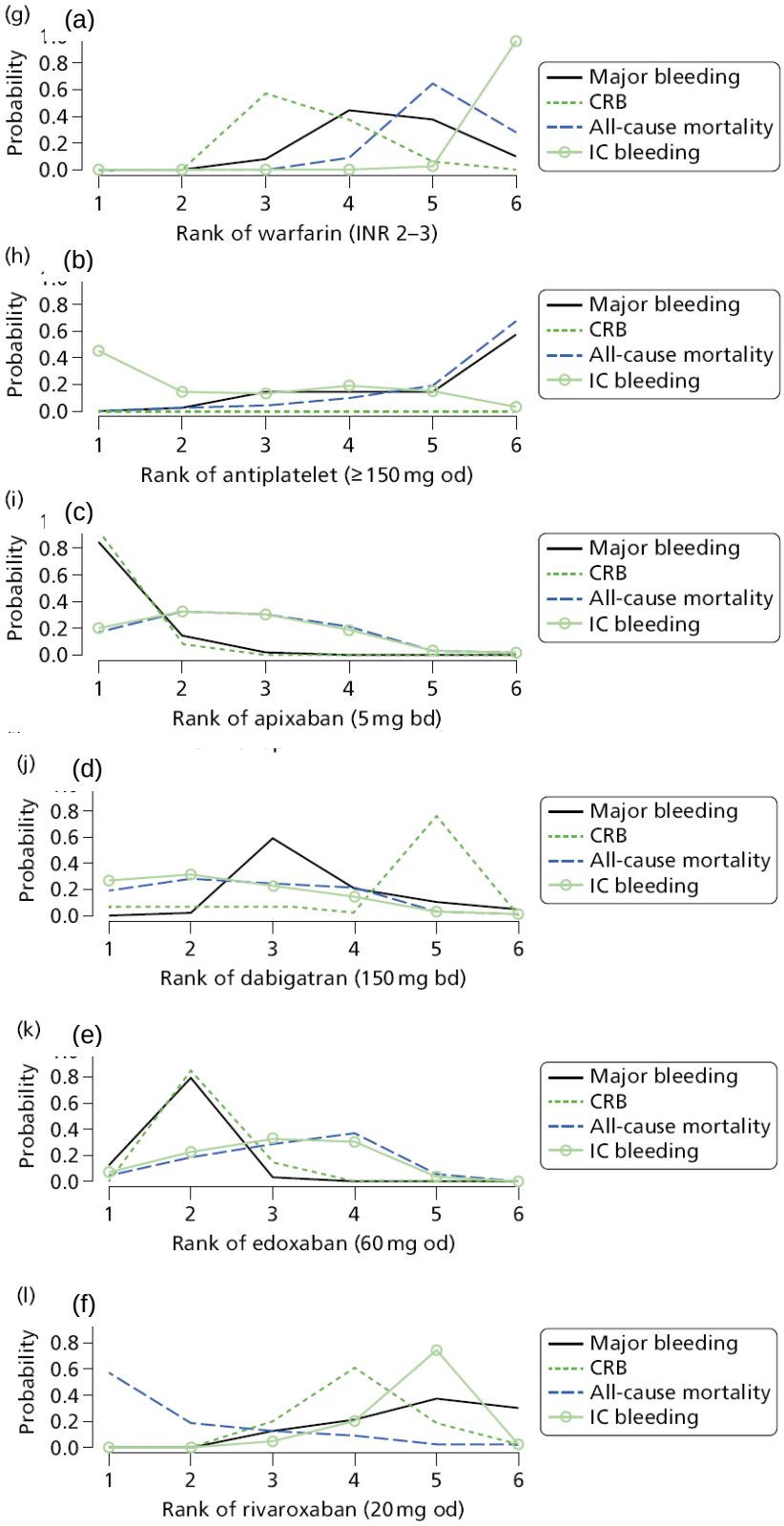


Figure 73: Ranking of safety (Rank-o-gram) for licensed interventions examined in stroke prevention in AF

Ranking indicates the probability to be the best treatment, the second best, the third best, and so on, among the NOACs (and warfarin). A higher probability (y-axis) at a rank of 1(x-axis) indicated the most favourable treatment for the plotted endpoint (major bleeding, CRB, all-cause mortality, or IC bleeding).

7.9 Findings (Cost-effectiveness analysis: Stroke prevention in AF)

7.9.1 Introduction

Further to section 7.7 where the protocol for the CEA is described, in this section the results of the CEA for first-line treatment of patients with AF are presented. This area of the research project was undertaken by the health economics team, with inputs for the model and sensitivity analyses informed by the clinical team members. Results are presented from Bayesian analyses with 95% credible [confidence] intervals.

7.9.2 Summary from NMA

Results from the NMA are summarised in Table 76.

Table 76: Mean HR and 95% Credible Intervals relative to warfarin from the NMA for each event and treatment (included in the CEA model)

NOAC	Ischaemic stroke	TIA	SE	ICH	Other CRB	MI	Death (all cause)
Apixaban (5mg bd)	0.90 (0.72 to 1.11)	0.74 (0.041 to 3.26)	0.65 (0.33 to 1.18)	0.46 (0.36 to 0.58)	0.82 (0.70 to 0.94)	0.86 (0.65 to 1.10)	0.89 (0.8 to 0.99)
Dabigatran (150mg bd)	0.75 (0.58 to 0.97)	2.68 (0.062 to 16.10)	0.65 (0.52 to 0.80)	0.36 (0.26 to 0.49)	1.07 (0.92 to 1.24)	1.27 (0.93 to 1.68)	0.88 (0.77 to 1.00)
Edoxaban (60mg daily)	1.00 (0.83 to 1.20)	2.76 (0.06 to 15.80)	0.58 (0.30 to 0.97)	0.49 (0.39 to 0.61)	0.88 (0.82 to 0.94)	0.95 (0.74 to 1.19)	0.92 (0.83 to 1.01)
Rivaroxaban (20mg daily)	0.92 (0.73 to 1.13)	2.68 (0.063 to 15.90)	0.95 (0.79 to 1.13)	0.65 (0.46 to 0.89)	1.05 (0.98 to 1.13)	0.79 (0.61 to 1.01)	0.83 (0.69 to 0.99)
Apixaban (2.5mg bd)	0.74 (0.042 to 3.37)	0.76 (0.041 to 3.51)	0.48 (0.031 to 1.97)	2.78 (0.06 to 16.2)	0.63 (0.080 to 2.06)	1.01 (0.049 to 4.67)	1.03 (0.050 to 5.03)
Dabigatran (110mg daily)	1.13 (0.89 to 1.42)	2.82 (0.062 to 16.40)	0.90 (0.73 to 1.10)	0.31 (0.22 to 0.43)	0.94 (0.81 to 1.09)	1.29 (0.94 to 1.71)	0.91 (0.80 to 1.03)

Table 77: Main assumptions in the AF model

	Assumption
1	Does not include minor non-clinically relevant bleeds as transient events
2	No distinction between severity of ischaemic strokes (SE assumed to be a transient event without long-term consequence)
3	Dose of Apixaban and dabigatran given does not reduce as patients age
4	Bleeds and ICH (and with it, haemorrhagic stroke) have the same effect on future risk of death as stroke
5	Patients switch to no treatment after ICH / haemorrhagic stroke
6	Patients on dabigatran who experience an MI will always switch to warfarin (not another NOAC)
7	Patients may switch from NOAC to warfarin or warfarin to no treatment after ischaemic stroke, bleed, SE or TIA
8	Patients may discontinue warfarin treatment or switch from a NOAC to warfarin, even if they do not experience an event (due to lack of compliance)
9	Warfarin arms from the RCTs identified in the systematic review are representative of the AF population in England and Wales
10	Events rate and relative treatment effects are assumed not to vary with age
11	Relative mortality rate in patients with AF relative to the general population does not vary with age
12	Warfarin treatment costs over 3 months are taken from the NICE costing report. Uncertainty in this is represented using a uniform distribution from 50% to 150% of the NICE costing report estimate
13	Assume no monitoring or administration costs for NOACs
14	Assume post-ICH management costs to be similar to post-ischaemic stroke management costs
15	Combined management costs for post-multiple event states (e.g. MI + stroke) to be the maximum of management costs for constituent events
16	Assume quality of life for patients with a history of multiple events to be a multiplicative combination of QoL for constituent events
17	All endpoints are equally weighted

7.9.3 Results of the cost-effectiveness model (AF)

Assumptions made in the construction of the AF model are listed in Table 77. Total costs and QALYs for first-line anticoagulation strategy are presented in Table 78. Incremental costs and QALYs for each strategy compared with the established standard of care, warfarin (INR 2-3) is also given.

Dabigatran (150mg twice daily) has the lowest expected total cost (£23,064), followed by apixaban (5 mg twice daily), edoxaban (60 mg once daily), warfarin (INR 2–3) and rivaroxaban (20 mg once daily), which has the highest expected total cost (£24,841). Expected costs are similar across all treatments, and there is a high degree of uncertainty around the costs for all treatments.

Apixaban has the highest expected QALYs (5.49), followed by rivaroxaban (5.45), dabigatran and edoxaban (both with 5.41), and warfarin (INR 2–3) (5.16). The NOACs have similar expected QALYs, all of which are higher than for warfarin (INR 2–3). There is a high degree of uncertainty around the QALY estimates.

At a willingness-to-pay threshold of £20,000 per QALY, all NOACs have positive expected incremental net benefit (INB) compared with warfarin (INR 2–3), suggesting that they may be a cost-effective use of NHS resources. Apixaban has the highest expected INB (£7,533), followed by dabigatran (£6,365), rivaroxaban (£5,279) and edoxaban (£5,212). Apixaban is the only NOAC for which the 95% CI around INB is positive, suggesting that apixaban is cost-effective compared with warfarin. These conclusions also hold at the higher threshold of £30,000. These assessments are made on non-discounted costs, using commercially available (i.e. non-commercially sensitive) prices. Based on experience that discounts for branded medicines within the same therapeutic class are generally similar, the above conclusions would continue to hold true.

The analysis was undertaken on the assumption that each of the endpoints were weighted equally, i.e. one ischaemic stroke was equal to one SE, although the economic implications of each event were

modelled using data from the ONS and NHS reference costs. The key drivers of the results favouring apixaban (5mg twice daily) were the lower rates of MI, ICH and other CRB (see Table 76). Although it is acknowledged that all-cause mortality could from a patients' perspective be regarded as the most important safety endpoint, it is possible that all-cause mortality is dominated by causes of death that are unrelated to the intervention so any benefit (or harm) will be concealed. Furthermore, the estimates generated from the CUA model carry a degree of uncertainty around the mean and 95% CrI values, which adds to the insensitivity and lack of specificity already present with the all-cause mortality outcome.

For apixaban (5mg twice daily), the estimates for TIA, MI, ICH and other CRB all fall in favour of this particular NOAC, with ischaemic stroke and all-cause mortality estimates broadly similar to the others. More specifically, the high cost and disutility of ICH has a great influence on total costs, total QALYs and net benefits. Apixaban has a low rate of TIA; however the minimal impact of this event means that it is not a driving factor in the results. Although dabigatran has the lowest rate of ICH, the higher rate of MI offsets this benefit.

Table 78: Cost-effectiveness of first-line treatment strategies for patients with AF

Estimated costs and outcomes	Warfarin (INR 2-3): mean (95% CI)	Apixaban (5mg bd): mean (95% CI)	Dabigatran (150mg bd): mean (95% CI)	Edoxaban (60mg daily): mean (95% CI)	Rivaroxaban (20mg daily): mean (95% CI)
Expected total costs (£)	24,418 (12,189 to 50,365)	23,340 (12,842 to 45,753)	23,064 (12,674 to 46,075)	23,985 (13,098 to 46,319)	24,841 (13,198 to 47,603)
Expected QALYs	5.17 (3.63 to 6.54)	5.49 (3.84 to 6.79)	5.41 (3.82 to 6.70)	5.41 (3.82 to 6.68)	5.45 (3.82 to 6.80)
Expected incremental total costs (£)	Reference	–1078 (–7626 to 2568)	–1354 (–8049 to 2273)	–433.4 (–6430 to 3619)	422.5 (–4730 to 5104)
Incremental expected QALYs	Reference	0.3227 (–0.01486 to 0.8142)	0.2505 (–0.08034 to 0.7025)	0.2389 (–0.1122 to 0.6841)	0.2851 (–0.06816 to 0.8096)
Incremental expected net benefit (£20,000)	Reference	7,533 (489.9 to 18,228)	6,365 (-167.7 to 17,039)	5,212 (-893.8 to 14,826)	5,279 (-1,097 to 15,180)
Incremental expected net benefit (£30,000)	Reference	10,760 (576.2 to 25,861)	8,871 (-597.3 to 23,402)	7,601 (-1,556 to 20,987)	8,130 (-1,399 to 22,819)

Incremental results are relative to warfarin (INR 2-3); Figures are presented as mean (CI)

7.9.4 Results of the sensitivity analyses

To explore whether or not results were sensitive to the assumed costs of warfarin, the extreme case in which there is no administration or monitoring costs for warfarin was reviewed first. This assumption found little effect on the conclusion that apixaban (5 mg twice daily) is the most cost-effective strategy (see Figure 74). The rationale for this exercise was to prove that if warfarin is not cost-effective with zero monitoring costs then it will not be cost-effective with monitoring costs greater than this. On this principle, sensitivity analyses with higher monitoring costs were not carried out. Similarly, the assumption that ICH and other CRBs have no effect on future mortality risk did not alter the conclusion that apixaban is most likely to be cost-effective (see Figure 75).

Different treatment switching strategies were also explored. If patients switch to no treatment only when they experience an ICH or a MI (if on dabigatran), the results are similar to the primary analysis. If all patients switch treatments after ischaemic stroke, bleed, SE and TIA, in addition to the switching after ICH and MI (for dabigatran) then patients spend only a short time on a NOAC before switching to warfarin. In this scenario, it is perhaps unsurprising that warfarin is the most cost-effective strategy. Another switching strategy was also considered by which all patients switch after an ischaemic stroke or clinically relevant bleed, and none switch after a TIA or SE, and found that the results are similar to the primary analysis.

For a comprehensive analysis, different initial ages (and corresponding doses) for the cohort were also explored. Apixaban (5 mg twice daily) is the strategy that is most likely to be cost-effective, assuming an initial cohort age of 60 years. Lower doses of apixaban (2.5 mg twice daily) and dabigatran (110 mg twice daily) are recommended for elderly patients, and were compared in a sensitivity analysis (see Figure 76). The uncertainty is much greater in this comparison, but apixaban (2.5 mg twice daily) is most likely to be the most cost-effective first-line therapy for the prevention of stroke in AF.

Although there was a high degree of uncertainty in the inputs of the model, apixaban (5 mg twice daily) was identified with the highest probability of being the most cost-effective first-line treatment over a range of willingness-to-pay-per-QALY thresholds. The driver of this result is the generally lower rates of MI, ICH and other CRB on apixaban than the other NOACs. The assumptions made within the model are summarised in Table 77. The conclusion remained robust to a range of sensitivity analyses. The only sensitivity analysis found to affect the conclusion was the assumption about treatment switching strategy; if treatment switching is assumed to always occur after stroke, bleed, SE or TIA then warfarin was identified as the most cost-effective treatment. However, based on discussion between the clinical group members it was agreed that this extreme switching strategy was not considered realistic in practice. The costs of warfarin were taken from the NICE costing report (304). As a level of uncertainty is described in this estimate it was agreed that an extreme case scenario analysis should be performed in which it is assumed that warfarin treatment and monitoring incurs zero cost. Under this assumption apixaban (5mg twice daily) was still the most cost-effective treatment. Apixaban and dabigatran may be given in lower doses to the elderly. It was assumed that all patients would receive the higher dose, and remain on it, even as they age, however, results were robust to a sensitivity analysis assuming only the lower doses of apixaban (2.5 mg twice daily) and dabigatran (110 mg twice daily) were administered.

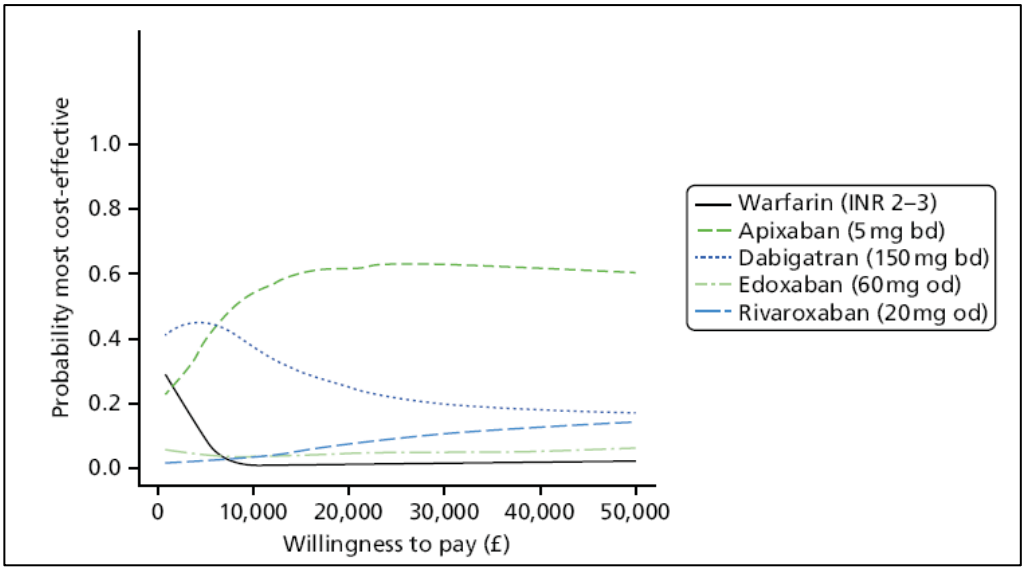


Figure 74: Cost-effectiveness acceptability curves (CEAC) for sensitivity analysis assuming that the cost of warfarin treatment is zero

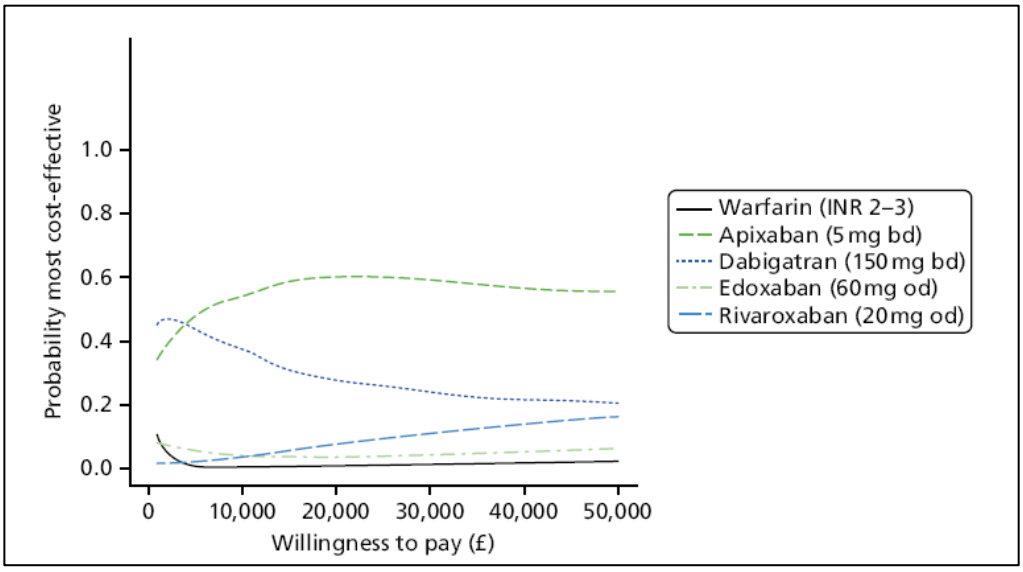


Figure 75: Cost-effectiveness acceptability curves (CEAC) for sensitivity analysis assuming no effect of bleed or ICH on mortality risk

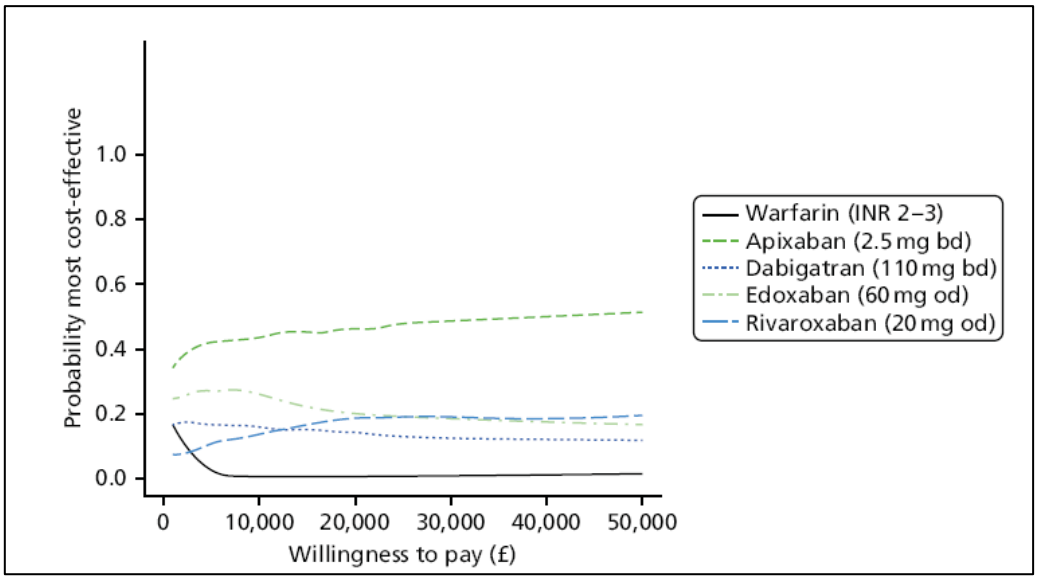


Figure 76: Cost-effectiveness acceptability curves (CEAC) for sensitivity analysis comparing lower doses of apixaban and dabigatran, as would be administered in older patients with AF

7.9.5 Limitations

It was not possible to include betrixaban within the analyses due to lack of evidence; therefore no conclusions are made about the relative cost-effectiveness of betrixaban or other unlicensed treatments. An assumption was made that age determines mortality rate, but that other event rates and relative treatment effects do not depend on age. The effects of minor and major stroke are not distinguished within the model. Some previous models have done so (256, 305, 306) but the health economics team found that there was insufficient evidence to be able to estimate rates differently. It was assumed that SE is a transient event with no long-term consequences. Although there can be long-term consequences, such as limb loss, these are very rare, and as such one would not expect inclusion of these to affect the results.

One notable potential limitation of the model is that there is no distinction between different types of AF. There is emerging evidence that there may be a 'dose–response' relationship in stroke risk with increasing 'persistence' of AF,(307) although others have suggested that risk of stroke is as high in paroxysmal patients with AF as with persistent or permanent AF.(308) The RCTs included in this review are likely to have recruited mostly persistent or permanent patients with AF, therefore the conclusions may not extend to patients with paroxysmal AF.

There have been few analyses of NOACs for the prevention of stroke in AF in the UK population. Kansal et al.(255) found dabigatran to be cost-effective compared with warfarin and aspirin in the UK setting, as was the case in the current analysis. However, the authors did not include any other NOACs. The Bayer submission to NICE on rivaroxaban(253) found it be cost-effective compared with warfarin. This submission also found rivaroxaban and dabigatran to have equivalent effects but dabigatran to have higher costs, thus concluding that rivaroxaban is the most cost-effective. Within that analysis only rivaroxaban was compared with warfarin, which found close to a 60% probability that rivaroxaban was cost-effective in the

£20,000–30,000 threshold range, similar to the current analysis probability that a NOAC (apixaban) was most cost-effective. The Harrington et al.(309) model, conducted in the US setting, compared apixaban (5 mg twice daily), dabigatran (110 mg twice daily), rivaroxaban (20 mg once daily), and warfarin (variable dose), and found that apixaban had the highest expected QALYs, followed by dabigatran, rivaroxaban and warfarin. This concurs with the current analysis which also found apixaban to have the highest expected QALYs and that dabigatran and rivaroxaban would have higher expected QALYs than warfarin, although the high degree of uncertainty in the current results renders them compatible with the order found by Harrington et al. These authors also found apixaban and dabigatran to be cost-effective compared with warfarin, and other US studies found apixaban,(310) rivaroxaban(305) and dabigatran(256) to be cost-effective compared with warfarin. Although costs in the USA are not strictly comparable with those in the UK setting, it is promising to see that the results from the current analysis are in line with these earlier findings.

7.10 Discussion and Conclusion

7.10.1 Clinical effectiveness

There was evidence that apixaban (5 mg twice daily), dabigatran (150 mg twice daily), edoxaban (60 mg once daily) and rivaroxaban (20 mg once daily) all reduce the risk of stroke or SE compared with warfarin (INR 2–3). Among the NOACs, there was evidence of a higher risk of stroke or SE with edoxaban and rivaroxaban than dabigatran.

There was evidence that dabigatran reduces the risk of ischaemic stroke compared with warfarin, whereas edoxaban increases that risk. There was little evidence that the risk of ischaemic stroke differed between licensed doses of NOACs.

There was weak evidence that the risk of MI is higher with dabigatran (110 mg twice daily), dabigatran (150 mg twice daily) and edoxaban (30 mg once daily) than warfarin (INR 2–3), and weak evidence that the risk of MI is lower with rivaroxaban (20 mg

once daily) than warfarin (INR 2–3). Among the NOACs, there was weak evidence that MI risk is higher with dabigatran (150 mg twice daily) than apixaban (5 mg twice daily), and lower with rivaroxaban (20 mg once daily) than dabigatran (150 mg twice daily).

There was evidence that apixaban (5 mg twice daily), dabigatran (110 mg twice daily), edoxaban (30 mg once daily) and edoxaban (60 mg once daily) all reduced risk of major bleeding compared with warfarin (INR 2–3). Among the NOACs, there was evidence that risk of major bleeding is higher with dabigatran (150 mg twice daily) than apixaban (5 mg twice daily), and with rivaroxaban (20 mg once daily) than apixaban (5 mg twice daily) and edoxaban (60 mg once daily).

There was evidence that the risk of CRB during antiplatelet therapy (aspirin < 150 mg once daily) is lower than with warfarin (INR 2–3). There was evidence that the risk of CRB with apixaban (5 mg twice daily), edoxaban (30 mg once daily) and edoxaban (60 mg once daily) is also lower than with warfarin (INR 2–3). However, edoxaban (30 mg twice daily) and edoxaban (60 mg twice daily) increased CRB compared with warfarin (INR 2–3). In comparisons among NOACs, there was evidence that CRB with edoxaban (60 mg once daily) and rivaroxaban (20 mg once daily) is higher than with apixaban (5 mg twice daily), and that rivaroxaban (20 mg once daily) increases CRB compared with edoxaban (60 mg once daily).

There was strong evidence that risk of intracranial bleeding was lower with apixaban (5 mg twice daily), dabigatran (110 mg twice daily), dabigatran (150 mg twice daily), edoxaban (30 mg once daily), edoxaban (60 mg once daily) and rivaroxaban (20 mg once daily) than warfarin (INR 2–3). For each of these NOACs [and doses], except for rivaroxaban (20 mg once daily), the estimated relative risk reduction for intracranial bleeding was > 50%. There was weak evidence that risk of intracranial bleeding is higher with rivaroxaban (20 mg once daily) than apixaban (5 mg twice daily), dabigatran (150 mg twice daily) and edoxaban (60 mg once daily).

Risk of all-cause mortality was lower with apixaban (5 mg twice daily), dabigatran (110 mg twice daily), dabigatran (150 mg twice daily), edoxaban (30 mg once daily), edoxaban (60 mg once daily) and rivaroxaban (20 mg once daily) than warfarin (INR 2–3), but there was little evidence of a difference between the licensed doses of NOACs for this outcome.

7.10.1.1 Summary – Clinical Trade-off Analysis

Overall, apixaban (5mg twice daily) was ranked as being among the best interventions for a wide range of the outcomes that were evaluated, including stroke or SE, MI, major bleeding, ICH and all-cause mortality. Edoxaban (60 mg once daily) was ranked second for major bleeding and all-cause mortality. Except for all-cause mortality, outcomes for rivaroxaban (20 mg once daily) were ranked less highly than several other NOACs. The non-NOAC interventions [warfarin (INR 2–3) and antiplatelet therapy (aspirin/clopidogrel ≥ 150 mg od)] were ranked worst for stroke or SE and were not among the best three interventions for any of the outcomes. Apixaban (2.5 mg twice daily) or betrixaban (40 mg once daily) were not included within the analyses because comparisons involving these interventions were imprecisely estimated.

In the sensitivity analyses, results were similar when using HRs instead of ORs. Moreover, no evidence of effect modification according to mean TTR for patients on warfarin was found. An important limitation is that primary studies did not report the mean time above or below therapeutic range for warfarin arms. Therefore, it was not possible to address some clinically relevant questions regarding the impact of treatment settings for warfarin on stroke prevention, as well as on bleeding and other AEs.

7.10.2 Cost effectiveness

Dabigatran (150 mg twice daily) has the lowest expected total cost (£23,064), followed by apixaban (5 mg twice daily), edoxaban (60 mg once daily), warfarin (INR 2–3) and rivaroxaban (20 mg once daily), which had the highest expected total cost (£24,841). Expected costs are similar across all treatments, and there is a high degree of uncertainty around the costs for all treatments.

Apixaban (5 mg twice daily) has the highest expected QALYs (5.49), followed by rivaroxaban (20 mg once daily) (5.45), dabigatran (150 mg twice daily) (5.42), edoxaban (60 mg once daily) (5.41) and warfarin (INR 2–3) (5.17). The NOACs have similar expected QALYs, all of which are higher than for warfarin (INR 2–3). There is a high degree of uncertainty around the QALY estimates.

At a willingness-to-pay threshold of £20,000 per QALY, all NOACs have positive expected incremental net benefit (INB) compared with warfarin (INR 2–3), suggesting that they may be a cost-effective use of NHS resources. Apixaban (5 mg twice daily) has the highest expected INB (£7533), followed by dabigatran (150 mg twice daily) (£6365), rivaroxaban (20 mg once daily) (£5279) and edoxaban (60 mg once daily) (£5212). Apixaban (5 mg twice daily) is the only NOAC for which the 95% CI around INB is positive, suggesting with confidence that apixaban is cost-effective compared with warfarin. These conclusions also hold at the higher threshold of £30,000. The key drivers of these results are the lower rates of MI, ICH and other CRB for apixaban (5 mg twice daily).

7.10.2.1 Summary – Cost-Effectiveness Trade-off Analysis

The CEAC indicates that apixaban (5 mg twice daily) has the highest probability of being the most cost-effective first-line therapy for AF, close to 60% in the £20,000–30,000 range of willingness-to-pay thresholds generally considered by NICE. Dabigatran (150 mg twice daily) has the highest probability of being cost-effective if the willingness-to-pay threshold is very low as a result of having the lowest expected total costs.

Warfarin (INR 2–3) and edoxaban (60 mg once daily) are unlikely to be cost-effective. Apixaban (5 mg twice daily) has the highest expected net benefit at a wide range of willingness-to-pay thresholds. In conclusion, apixaban (5 mg twice daily) is likely to be the most cost-effective first-line therapy for AF, under the assumptions of the model.

7.10.3 Analysis of the value of information for future research

The optimal decision regarding the most cost-effective NOAC is most sensitive to the HRs comparing the NOACs, suggesting that a head-to-head trial comparing NOACs is of value. The decision is also sensitive to costs, the effect of past events on future HRs, and probabilities of treatment switching. A head-to-head trial could also provide information about baseline event rates, costs and switching probabilities. However, a study powered to measure all of these outcomes with sufficient precision would require a very large sample size, which may be prohibitively expensive. With these agents reaching the end of their patent protected period (dabigatran potentially available as a generic medicine as early as 2018) it is highly unlikely that this is of commercial interest to any of the Pharmaceutical companies who have still yet to deliver peak sales.

7.10.4 Strengths

The strengths of this appraisal, undertaken as an NIHR Health Technology Assessment, includes its comprehensive coverage of all the therapeutic areas in which NOACs have been evaluated to date, using the same methodology [only AF reported here]. Previous analyses of comparative effectiveness have focused on individual therapeutic areas, making it more difficult to judge if one of the four licensed NOACs might emerge as a frontrunner in more than one therapeutic area.

Additional strengths include careful appraisal of study quality; focus on clinically relevant end points; an evaluation of safety that considers evidence spanning all therapeutic areas together, to maximise power; the development of a treatment hierarchy for the

different anticoagulant indications, where the data allowed it; and a CEA that is relevant to the NHS. This was possible through setting up a HTA team that encompassed academics from the systematic review and health economics specialties as well as clinical and patient partners.

7.10.5 Limitations

The limitations of this NIHR Health Technology Assessment, as is the case for systematic reviews and meta-analysis, relate to shortfalls in the primary data on which the overview is based. These include:

- No direct head-to-head comparisons between different NOAC drugs – all such comparisons were therefore based on indirect evidence derived from the networks
- Economic analyses for long-term conditions such as AF make long-term projections on the basis of short-term trial evidence, observational data and clinically informed assumptions about plausible treatment pathways and health-state transitions. These assumptions and evidence limitations are discussed above.
- The profile of patients entering trials may not be the same as those treated in practice, who may be older and have more comorbidities. Treatment benefits in such patients may be smaller, and rates of harm higher, than estimated by trials.
- As for all new drugs, adverse effects that remained undetected during development may come to light with high-volume use post licensing.
- It is possible that patients treated with warfarin in practice are at higher risk of bleeding complications than those in trials because of a greater number of comorbidities and less stringent control of anticoagulation. However, concerns have also been raised previously that the time spent in the therapeutic range was suboptimal among patients in clinical trials who were assigned to warfarin. Thus, clinical trials could have underestimated both the benefits and the risks of warfarin treatment.

Several factors led to imprecision in the estimation of certain treatment effects. These included low rates of occurrence of certain end points; widespread use of composite end points, with low rates of occurrence of certain (more clinically relevant) components of the composite; as well as substantial inconsistency in the reporting of end points in different trials in the same therapeutic area, leading to a substantial number of missing end point data (see Table 39 and Table 40). Owing to the low event rates and lack of substantial replication of specific comparisons across studies, it was agreed that a fixed-effects models should be used for the NMA. This does not account for heterogeneity in treatment effects. Under fixed-effects models, the Bayesian analyses with vague priors will produce results very similar to frequentist analyses.

The apparent efficacy of NOACs when compared with warfarin could be inflated if control of the INR was suboptimal among patients who were randomised to warfarin. For this reason, many of the studies reported time spent in the therapeutic range (TTR), as an index of anticoagulant control. This is a potentially important issue for the studies of stroke prevention in AF, for which 16 (73%) of the 22 studies that included a warfarin intervention arm reported mean TTR. There was substantial variation in TTR (from 45.1% to 83%) between these studies. The pre-specified protocol for this health technology appraisal specified TTR as a potential modifier of NOAC treatment effect in trials in which warfarin was the comparator.

In some situations, the need for anticoagulation monitoring with warfarin treatment may be viewed as a useful means to confirm adherence to anticoagulant therapy rather than as an inconvenience. This was an interesting point raised by the patient partner on the panel who suggested that an awareness of the activity of the medicine (i.e. INR testing) provides positive reassurance that the treatment is working and will reinforce patients taking their medicine, compared with NOACs which do not require routine testing or monitoring.

There is a suggestion that the efficacy and safety of dabigatran could be improved by monitoring of achieved drug levels, like with warfarin, because these exhibit wide inter-individual variation. This will reduce the convenience advantages of this NOAC and increase its cost compared with warfarin or other NOACs, however as this is not a requirement this was not modelled in the cost analyses. Only one of the studies included within the current review considered whether or not monitoring improves the efficacy and safety of NOACs: in a subsample of 9,183 patients in the RE-LY trial,(245) ischaemic stroke and major bleeding both correlated with dabigatran plasma concentrations. Specific tests to measure the anticoagulation effects of NOACs are being developed but are not yet widely available(311) and routine coagulation tests such as prothrombin time and activated partial thromboplastin time are of limited use.(312, 313) It is therefore currently unclear whether or not the efficacy and safety profiles of NOACs can be improved by monitoring and dose adjustment. Monitoring may be particularly helpful in certain clinical situations (for example, in the peri-operative period around planned or emergency surgery, or patients presenting with bleeding(313)) and in certain patients groups (for example, advanced age and renal impairment [outside of the range investigated by the clinical trials]).

Finally, therapeutic decision making may be influenced by recognition that effective treatments for reversal of anticoagulation with NOACs are still under development. Promising results from early phase trials of medicines currently identified include:

- Aripazine (PER-977; PER 977; ciraparantag) is a synthetic cationic molecule that binds unfractionated and LMWH, the factor Xa inhibitors edoxaban, rivaroxaban and apixaban, and the factor II inhibitor dabigatran, but not to warfarin.(314) In a Phase I trial involving 80 healthy volunteers, intravenous PER977 reversed the prolongation of whole blood clotting time induced by a single oral dose of edoxaban 60 mg in a dose-dependent fashion, within 10–30 minutes of administration.238 Phase II clinical studies of this agent are in progress.

- Andexanet alpha (PRT4445; PRT064445) is a recombinant modified factor Xa molecule that acts as an antidote to factor Xa inhibitors through a decoy mechanism. A number of Phase III studies of this agent are under way.(315, 316)
- Idarucizumab (BI 655075) is a humanised monoclonal antibody fragment that binds dabigatran to reverse its anticoagulant activity.(317, 318) Phase I/II studies of this agent have been completed. A Phase III study investigating reversal of anticoagulation in patients receiving dabigatran who have uncontrolled bleeding or who require emergency surgery or invasive procedures is under way.

7.10.6 Future research

Like most NMA of new therapies, the evidence on the comparative efficacy and safety of NOACs in this review has come exclusively from indirect comparisons; this again is a consequence in the absence of regulatory requirement to undertake head-to-head trials, although is more likely to be attributable to this new group of anticoagulants undergoing trial investigation at similar times before each has a chance to become standard practice and therefore a comparator. A mechanism by which to compare them whilst under development would have been useful to serve as a guide at the time of regulatory approval and launch. Among patients with AF, a long-term condition, the trials have also been of relatively short duration. As different manufacturers have developed each of the agents evaluated in this review on the premise that it will achieve a healthy market share and therefore financial return, it is highly unlikely that any head-to-head trials will be initiated by the industry for academic gain.

Reliable estimation of the cost-effectiveness of NOACs in different clinical scenarios requires high-quality data on absolute event rates for the various efficacy and safety outcomes. NHS health record data could provide a rich source for information, but so far health record data have been insufficiently utilised for this purpose. As NHS Digital evolves within the UK this may be a possibility in the future.

Although NOACs were developed in part to supersede warfarin by obviating the need for therapeutic monitoring of anticoagulation, to improve convenience, recent studies have suggested that monitoring of drug levels may improve safety and efficacy of dabigatran treatment. Whether or not this is also the case for other NOACs is not known and is of course associated with an additional cost not factored within this analysis.

The requirement for therapeutic drug monitoring with warfarin also serves as a means to assess adherence. Thus far, long-term adherence rates for NOACs (e.g. among patients with AF who may require anticoagulation for many years) have not been evaluated.

The research needs identified by this review are therefore as follows:

- To undertake one or more trials making direct comparisons between the most promising NOACs and NOAC doses, in situations typical of NHS clinical practice (potentially apixaban vs. rivaroxaban)
- Information on long-term rates of the main efficacy and safety outcomes among patients receiving anticoagulants for AF (e.g. from registries or health record data)
- Information on the role (if any) of therapeutic monitoring to enhance the safety and efficacy of NOACs
- Information on long-term adherence rates in patients receiving NOACs for AF

7.10.7 Summary

This systematic review, Bayesian NMA, clinical trade-off analysis and CUA, provides the most comprehensive and up-to-date analysis of the benefits, safety issues, and cost implications for NOACs compared with warfarin for the prevention of stroke in AF. At licensed doses and at NHS prices, a number of NOACs are of net benefit compared with warfarin with apixaban being ranked the highest. A trial directly comparing NOACs would overcome the need for indirect comparisons to be made using NMA, however

this is unlikely to occur. Based on the NMA and CEA findings, apixaban emerged as the best intervention. This recommendation is supported by a registry study which found that patients treated with apixaban were associated with a lower risk of major or CRNM bleeding compared with warfarin.(319)

This work has been published by the NIHR as part of their Health Technology Assessment programme: Volume 21, Issue 9 (March 2017) and meets the criteria for inclusion on the NIHR Centre for Reviews and Dissemination (CRD) Register.

7.10.8 Implications for clinical practice

The findings from the NMA and CEA broadly suggest that all NOACs dominate warfarin for the prevention of stroke in AF. Overall, apixaban appears to be the most cost-effective NOAC owing to the highest expected QALY and greater confidence of the 95% CI around the incremental net benefit estimate. However, the absolute values between NOACs are unlikely to be so significant so as to actively modify therapy for patients currently on treatment. Local guidelines will need to be drawn up in conjunction with clinical specialists and patient partners to determine the benefits of one NOAC over another in determining the favoured agent from the class.

[BLANK PAGE]

8.0 Chapter 8: Discussion

8.1 Summary

The application of a quantitative trade-off analysis, including efficacy, safety, tolerability, and / or cost parameters, as a supplement to Frequentist or Bayesian meta-analysis, over a number of clinical specialities has successfully aided in the establishment of a hierarchy of treatments. This hierarchy has been used to change practice within a pilot site of the UK NHS resulting in cost savings or cost-avoidance without compromising on patient care. Indirect comparisons, generated through NMA as conducted across a number of clinical specialities within this thesis, not only help develop rational treatment hierarchies, but also inform on the choice of high priority comparator agents for direct head-to-head analysis in the form of an RCT.

Findings from the quantitative trade-off analysis of antiepileptic drugs for refractory epilepsy and of antimuscarinics for overactive bladder syndrome have modified prescribing trends within North Central London, providing moderate cost avoidance and cost savings for the NHS. The findings from the meta-analysis and supporting cost-utility analysis of angiotensin receptor blockers for hypertension and heart failure have directed change in prescribing across the UK, providing a saving to the health economy of approximately £200 million without compromising patient safety. Lastly, the findings from the network meta-analysis and supporting cost-utility analysis of novel oral anticoagulants for prevention of stroke in AF are in the process of being circulated to support and standardise prescribing.

8.2 General study findings

The approval of new medicines currently entering the UK (and European) market is required to satisfy regulatory criteria. A selection of these are subject to additional commissioning assessment. Where more than one medicine from the same class is available and / or recommended this presents the clinician with a dilemma as to which to choose. This predicament affects the health

economy where choices are favoured towards newer, more expensive medicines on the basis that they must be more efficacious or better tolerated.

The consideration of many treatments within the context of a meta-analysis, in the absence of a head-to-head RCT, is becoming increasingly popular as clinicians and policy makers attempt to consider together and understand the results from multiple trials on the same topic. Although assumptions and considerations for such analyses are required from the clinical community, the methodology and execution of NMA were traditionally held in the realms of statisticians owing to the specialised software required. With the availability of WinBUGS, a publically available software, and clear technical guides, this may now be utilised by policy developers directly.

The studies presented within this thesis provide a concept of clarification for clinicians in practice regarding the most appropriate use of medicines available on balance of safety and efficacy, and cost where analysed. To explore its applicability, a number of different medicines across a range of different conditions were selected on the basis of their clinical importance. This includes treatments for refractory epilepsy (partial seizures with secondary generalisation), hypertension, heart failure, overactive bladder syndrome, and atrial fibrillation.

Across a range of specialities, the analyses within this thesis demonstrate that the data required for regulatory purposes are not sufficiently meaningful for clinical practice. The use of frequentist or Bayesian meta-analysis, with subsequent trade-off and / or cost-effectiveness analysis, are able to produce clinically-relevant data to inform clinicians and prescribing committees responsible for the development of guidelines. Where network meta-analyses produce findings that suggest no significant difference between the medicines under review, it can be reasonably assumed that the clinically-relevant difference between such agents is also not significant. Under such assumption a trade-off analysis can be supported by a cost-minimisation analysis. Where network meta-

analyses produce findings that are statistically significantly different, the trade-off analysis should be supported by a cost-utility analysis. These analyses can be used prospectively i.e. to prevent the unnecessary uptake of new medicines that do not appear to offer any clear added value over existing therapies, or retrospectively i.e. to disinvest in medicines currently available on hospital or general practice formularies owing to their low clinical value.

Among the topics considered within this thesis, the findings from the trade-off analysis for Chapter 5: Research Project Two – Angiotensin-II Receptor Antagonists for the Management of Hypertension and Heart Failure, and for Chapter 6: Research Project Three – Antimuscarinics and newer agents for the Management of Overactive Bladder syndrome have been converted into meaningful guidance for clinicians, via the relevant prescribing committee. This has subsequently led to changes in prescribing practice and savings to the health economy without compromises in patient care. The findings from the trade-off analysis for Chapter 4: Research Project One – Antiepileptic drugs for Refractory Epilepsy were implemented differently owing to the complexity of the condition. The most practical application of the results was agreed as being a benchmark for future medicines licensed for the same indication thereby restricting formulary adoption pending clear benefits on efficacy, tolerability, or cost. Within the local economy, North Central London, this has prevented widespread use of newer agents brought to market since the publication of this work. The work undertaken within the last research project, Chapter 7: Research Project Four – Oral anticoagulants for prevention of stroke in atrial fibrillation, was more extensive for a comprehensive cost-utility analysis to complement the trade-off analysis. Whilst changes in practice have yet to be implemented, the findings have been shared with the commissioners within North Central London and are being used to support the CCGs on the appropriate use of the newer NOACs compared with warfarin.

8.3 General study strengths and limitations

Meta-analysis of data has been used to synthesise results for several decades. Interest in this methodology continues as it enables researchers to reconcile inconsistent or unclear findings from individual RCTs and reach a definitive answer to the research question of interest. As a result, meta-analysis can overcome a key limitation of small studies, lack of sufficient statistical power.

The use of Bayesian network meta-analysis supplemented by trade-off and, where applicable, cost-effectiveness analysis employed during this thesis extends the above principle in order to further clarify research questions still faced in clinical practice. This methodology has several strengths, although some important limitations are apparent, as outlined within each chapter. The general strengths and limitations in terms of data and study validity are discussed below.

8.3.1 Data validity

A reservation amongst the clinical community when using meta-analysis techniques to infer treatment pathways is the quality of the primary data included. It is thus important to validate this step in accordance with robust processes.

To mitigate these concerns, strict guidelines are available for authors undertaking systematic review and meta-analysis. The use of an *a priori* protocol and data extraction form ensure that data used are consistently obtained. Following the principles outlined within PRISMA, the integrity of the data is maintained through capture of the appropriate patient groups, including those lost to follow-up or excluded from treatment for the efficacy and safety endpoints of interest.(320) To maintain relevance to the original scope, i.e. applicability to the UK NHS, the data considered within each of the review areas are applicable to UK licensed medicines and doses. Additionally, the restriction of full-text articles only ensures that the validity of the data has been assessed by an impartial and independent editorial board.

It is important to highlight that the work undertaken within this thesis has been peer reviewed, with three of the four areas published in medical journals, indicating the validity of the research technique and resultant findings. The findings have also been used to develop guidelines and implemented in practice indicating the acceptance of this technique amongst the relevant clinical communities.

8.3.2 Model validity

For the Frequentist analyses, the commercially available StatsDirect® software was used, utilising fixed-effect and random-effects models as appropriate, and standard reporting functionality. For the Bayesian analyses, the freely available WinBUGS® package was used. Coding for the Bayesian model were adapted from the NICE DSU Technical Service Documents and verified by experts in Bayesian analyses as described with the respective chapters. To demonstrate robustness in the Bayesian coding, results from Frequentist analyses were also presented to highlight the degree of correlation between the direct comparisons. As this is commonplace in the published literature it was replicated within this thesis. The aim of this was to assure the reader of the strength in the model when interpreting the indirect comparison estimates generated.

To further assure on the estimates generated, a range of assessments were undertaken as described in section 3.5.7 (including funnel plot, l'Abbé plot, and Higgins I-squared) and section 3.6.8 (including convergence, inconsistency and goodness-of-fit).

8.3.3 Bayesian vs. Frequentist Model

Bayesian and Frequentist methods both have advantages and disadvantages, with many similarities. The Frequentist analysis is used where data are available to enable direct comparisons. Where a framework is required to create indirect estimates the Bayesian analysis is used. When the sample size is large, Bayesian inferences often produce results for parametric models that are very similar to results produced by Frequentist methods. For all analyses conducted

within this thesis the sample sizes have been sufficiently large and uniform with a high degree of correlation between both methods.

Specific advantages in the use of Bayesian analysis include:

- Providing a principled way of combining prior information with new data within a solid decision framework (model). Previous information acquired is used to form a prior distribution for future analysis. All inferences logically follow Bayes' theorem.
- Providing exact inferences that are conditional on the dataset.
- Providing interpretable estimates, i.e. the true parameter has a probability of 0.95 of falling in a 95% credible interval
- Providing a convenient setting for a wide range of models, such as hierarchical (rank) models and missing data problems (indirect comparisons).

Specific disadvantages in the use of Bayesian analysis include:

- It does not advise users on how to select a prior, and there is no correct way to choose / use a prior. The Bayesian model therefore requires the user to possess skills in translating prior beliefs into a mathematically formulated prior. If a prior is used incorrectly, the results generated from the analysis may be misleading.
- As posterior distributions are influenced by the use of priors, it may be difficult to obtain acceptance of the results from subject matter experts who do not agree with the validity of the chosen prior.
- The analysis utilises a high degree of computational processing, particularly where the model includes a large number of parameters. The number of simulation / iterations / burn-in used may provide slightly different estimates, although this is less likely to be an issue where higher numbers (>10,000) are used.

8.3.4 Alternative Methods

All analyses undertaken within this thesis used a fixed- or random-effect model at the level of trials. The use of individual patient level data (IPD) instead of aggregate trial-level data in the execution of meta-analysis is an alternative methodology that could have been used. The use of aggregate data, defined as study level data obtained from study publications (or study authors), is traditional used when populating a dataset for meta-analysis. The use of IPD involves acquiring the raw individual level data for each study and used in the synthesis of the meta-analysis. In this technique, the same approach is used regarding performing the Frequentist or Bayesian meta-analysis however the preceding systematic review includes the additional step of drawing data from the authors original files rather than the summary data presented within a publication. It is proposed that meta-analysis of IPD has many potential advantages, both statistically and clinically, over meta-analysis of aggregate data.(321) Key examples include: aggregate data not available, poorly reported, derived and presented differently across studies, and more likely to be published where statistically or clinically significant. IPD data may also contain more participants and outcomes than that reported, and have a longer follow-up time.

IPD meta-analysis is not however without its disadvantages. In particular, this approach is resource intensive due to the substantial time (and costs) involved in contacting study authors, who then must be willing to part with their raw data. It would also be expected that dialogue with the author is required to clarify and resolve data issues. Additionally, there are concerns surrounding the applicability of ethics approval as handing of IPD may be akin to source data. Further, despite the best efforts of an author undertaking IPD meta-analysis, it may be possible that IPD is not available for all studies included. In this scenario an assessment will need to be undertaken to determine if the absence of IPD and use of aggregate data would bias the analysis. Lastly, it is important to acknowledge that the quality of IPD or aggregate data is dependent on the quality of the study itself.

The use of IPD meta-analysis was initially considered as part of the NIHR funded project (Chapter 7: Research Project Four – Oral anticoagulants for prevention of stroke in atrial fibrillation), however not followed through on balance of the perceived advantages compared with the complexity this technique would have added to the project.(322)

8.4 Study novelty

This research incorporates many novel features. First, this thesis describes a programme of studies that are of direct relevance to the health economy specifically related to medicines use. Second, the research undertaken within each chapter of the thesis describes the use of trade-off analysis using key efficacy and safety / tolerability end points based on clinical experience and implications for a hierarchy. Third, cost-effectiveness analysis are utilised (specifically cost-utility) which have been populated using UK relevant parameters and its outputs directly implemented within the NHS with consequent practice change.

During the period over which this thesis was conducted, a handful of leading statistical methodology teams have also recently identified the value of Bayesian network meta-analysis supplemented by a rank assessment. Examples include a review of first- and second-generation antidepressants,(323, 324) which would be similar to the review of AEDs undertaken in Chapter 4.0.

8.5 Reflective Account

The conduct of the various research projects within this thesis has provided me with the opportunity to translate many academic and clinical skills into practice. Over the course of the four research topics the complexity of the methodology increased requiring revision of the techniques both in coding and execution. The use of WinBUGS and network meta-analysis as a whole was a new experience; however its value has been accepted within the local medicines management committees once translated into clinically-relevant outputs such as guidelines or treatment hierarchies, made possible with the trade-off

technique. The ability to derive more robust, clinically useful information from pooling studies allowed me to see how clinical practice can be influenced and updated. The thesis as a whole has provided me with an aptitude for bridging the gap between regulatory data and data required at the interface of clinical practice.

The core principle of rankings from an NMA and their extrapolation to guide practice has been questioned recently in an article.(325) Although related, the trade-off technique presented within this thesis is not intended to generate model outputs to be used in an unqualified application. Conversely, the basis for establishing a rank follows request from clinical practice for clarity on specific parameters of interest in order to minimise uncertainty of the output. The result of this key difference being successful implementation of the findings, where applied.

8.6 Further Work

Numerous research projects could follow-on from this thesis. For example, the creation of a central database from published studies would be appropriate to enable meta-analysis (of a direct or indirect nature) upon the licensing of a new medicine within the same class. This could be automatically populated as part of the article submission process at the point of acceptance. The upload of data should be encouraged to include stratified and sub-group level data to enable sensitivity analysis if needed. The funding of such a system and requirement for journals to make their data available as a direct feed to this could be enforced as part of a legislative change in the regulatory process, for example an FDA and / or EMA requirement upon granting of the Marketing Authorisation. This database would be akin to the CPRD (Clinical Practice Research Datalink) currently established within the UK as a joint venture between the NIHR and the MHRA in order to improve drug safety, best practice and clinical guidelines. The CPRD is an example of how post-marketing surveillance programs are predominantly limited to safety monitoring, with the exception of those that are marketed under specific early access schemes.

The benefit of the above is that it would enable cumulative analysis to take place with data available in a standardised format, without the need for individuals to assess studies in detail to determine reporting as intention-to-treat, modified intention-to-treat, or per-protocol. This system would introduce a culture change in the contribution of cumulative analyses related to efficacy and place in therapy post-marketing. The execution of systematic reviews and meta-analysis should continue to be undertaken by experts within this area of research to ensure rigour is applied to the methodology and its appropriateness. The integrity of data will also need to be validated although there is potential that this could be done centrally if commissioned by the NIHR.

For certain medicines, the above methodology, Bayesian mixed treatment comparisons with trade-off analysis, is able to support the development of a rank order of medicines included within the scope, which in turn will describe the best agent based on efficacy and tolerability, second best agent etc. This may compound an issue already in existence; that the launch of certain medicines in clinical practice results in different patterns of prescribing compared with that examined and assessed within the clinical trial. The knowledge of a treatment being ranked, for example, third would from a guideline development perspective suggest that this should be positioned as third line, should a third choice be required. It is important to note that the trial data has not evaluated the use of these agents in sequential use and therefore the rank analysis findings are being extrapolated to infer that a third choice would deliver the same benefits and / or risks as being given third-line.

To answer this question a quite separate research project of n-of-1 studies would need to be raised. Having sought clarification from the MHRA, this would not constitute a clinical trial of an investigational medicinal product (CTIMP) as the medicines themselves would be used in accordance with their Marketing Authorisation with the trial being intra-subject. Given that this, like the projects undertaken within this thesis, is an academic rather than commercial exercise, in order to undertake this it would require administrative support from medicines management teams within Acute Trusts. To undertake this

formally, a potential expense lies on the availability of placebo- or active-matched stock in situations where subjective measures are the primary outcome. Such work on this is already underway at the National Hospital for Neurology and Neurosurgery (NHNN) where access to the new antiepileptic drug brivaracetam is being supplied on a named patient evaluation basis to patients in whom levetiracetam has previously been used but causing tolerability issues. On the basis of overlapping clinical pharmacology of these two agents, specific inclusion criteria set for the evaluation require a clinical response to levetiracetam and recording of off-target adverse effect(s), with outcomes evaluation for 6 months. A similar n-of-1 study is being discussed for eslicarbazepine secondary to oxcarbazepine.

8.7 Conclusion

While the randomised double-blind controlled clinical trial is the undisputed gold-standard for assessing drug efficacy, such trials are expensive to commission, include a fixed number of comparators, and usually do not provide the necessary comparisons to the potential standard of care. Moreover, they include limited patient numbers and run for short periods leading to limited power and therefore generalisability of its findings. Although designed to address statistical power as part of the calculation of recruitment target, the implications for clinical practice may be considered insufficient with specific issues related to comparative efficacy and safety. Therefore, at the time of product launch, clinicians are often faced with limited data and relevance for any new medicine. Despite numerous advances in evidence-based medicine, this continues to be an on-going gap.

Critical appraisal of systematic reviews and meta-analyses are embedded within evidence based medicine at the top of the pyramid on the basis that they can provide important insight into the utility of the results to improve patient care. Provided the methodology is robust in determining the validity, magnitude and applicability of the primary studies in accordance with the Cochrane Collaboration tools (or equivalent), the results will be generated to be of clinical value.

Meta-analyses frequently appear in the appraisal processes used by commissioning agencies, such as NICE, however they are not a mandatory part of the regulatory process. The research projects undertaken within this thesis demonstrate the merits of Bayesian mixed treatment comparisons, and the potential that Bayesian and frequentist meta-analysis can be enhanced to develop trade-off and cost-utility analyses. Routine implementation of such analyses can help to avoid uncertainty at the point of prescribing and unnecessary spend on newer medicines with no benefit over cheaper, established alternatives.

Although such system has great potential it will require investment and buy-in at the regulatory and research authority level internationally. It is however important to acknowledge that despite notable success, this system is subject to several limitation. First, individual prospective studies only provide information over a short period of time whilst the implications are extrapolated for life-long therapy. Second, excluding trials that have not been published may exaggerate the treatment effects observed. Third, although heterogeneity assessments are undertaken, pooling of data from trials with differences in trial design, methodology and patient groups may result in a heterogeneous dataset from which conclusions are drawn. However, such differences may serve to strengthen the meta-analysis by allowing generalisability of the results to a broader group as will be seen in general practice outside of the confines of a clinical trial. Fourth, analysis of this type will be restricted to the data available e.g. adult population; therefore any findings will need to be mindful that they are not extrapolated to patients beyond certain criteria. Fifth, doses employed within the clinical trials are not always consistent with those used in clinical practice, including the use (or absence) of a dose-escalation phase and close monitoring for dose-titration; the external validity of data may therefore be questionable, although it is sufficient to satisfy regulatory authorities. Finally, data within such analyses only relate to the use of the medicine under investigation for a specific place in therapy, generally first- or second-line; the generalisability or extrapolation of data beyond this should be considered with caution.

In this thesis I investigated the utility of a trade-off analysis as an extension to meta-analysis to overcome some of the limitations in practice currently. Data available publically from published articles, as accessed via on-line database and journals, was used to quantify the efficacy and safety / tolerability of medicines which are licensed for the same indications. These data were then converted into odds ratio, hazard ratio and / or number needed to treat or harm for specific endpoints considered of clinical importance in order to conduct a trade-off analysis and determine a rank. Provided a system is created centrally that captures the primary data in an organised manner, this methodology can be effectively utilised to inform on the relative merits and harms of every new medicine as part of product launch, thereby indicating potential value over existing treatments. The building of data from historical publications could be executed as a programme of responding under the guidance of the Regional Medicines Optimisation Committees (RMOC) as part of implementation of quality, standardised care across the UK.

Despite certain limitations with this type of analysis, something which is widely acknowledged with meta-analysis, given the efficiency and versatility of this method this could become a future standard following regulatory approval. This could facilitate a paradigm shift from the traditional model of trial-and-error, where new medicines are perceived to be better than existing medicines and receive uptake on this basis, to one of an informed rank. Improvement in prescribing trends and reductions in spend are crucial in order to achieve a sustainable National Health Service.

[BLANK PAGE]

9.0 References

1. NHS Digital. Prescribing Costs in Hospital and the Community2016. Available from: http://www.content.digital.nhs.uk/catalogue/PUB22302/hosp-pres-eng-201516-report.pdf.
2. Office for National Statistics. Expenditure on Healthcare in the UK: 20132015. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthcaresystem/articles/expenditureonhealthcareintheuk/2015-03-26/pdf.
3. O'Connor AB. Building comparative efficacy and tolerability into the FDA approval process. JAMA. 2010;303(10):979-80.
4. Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. BMJ. 2003;326(7400):1167-70.
5. Rohss K, Hasselgren G, Hedenstrom H. Effect of esomeprazole 40 mg vs omeprazole 40 mg on 24-hour intragastric pH in patients with symptoms of gastroesophageal reflux disease. Dig Dis Sci. 2002;47(5):954-8.
6. Jorgensen AW, Hilden J, Gotzsche PC. Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review. BMJ. 2006;333(7572):782. Epub 2006 Oct 6.
7. Sismondo S. Ghost management: how much of the medical literature is shaped behind the scenes by the pharmaceutical industry? PLoS Med. 2007;4(9):e286.
8. Sismondo S. How pharmaceutical industry funding affects trial outcomes: causal structures and responses. Soc Sci Med. 2008;66(9):1909-14. doi: 10.016/j.socscimed.2008.01.010. Epub Mar 4.
9. Lee K, Bacchetti P, Sim I. Publication of clinical trials supporting successful new drug applications: a literature analysis. PLoS Med. 2008;5(9):e191. doi: 10.1371/journal.pmed.0050191.
10. Angell M. The pharmaceutical industry--to whom is it accountable? N Engl J Med. 2000;342(25):1902-4.
11. Rogers A. European Parliament approves pharma law overhaul. Lancet. 2002;360(9343):1397-8.
12. Information CIfH. Drug Expenditure in Canada, 1985 to 2006. 2007. Available from: http://www.publications.gc.ca/site/eng/370361/publication.html.
13. (CADTH) CAoDaTiH. CADTH Presentation on the Common Drug Review to the House of Commons Standing Committee on Health2007.
14. Duerden M, Gogna N, Godman B, Eden K, Mallinson M, Sullivan N. Current national initiatives and policies to control drug costs in Europe: UK perspective. J Ambul Care Manage. 2004;27(2):132-8.
15. (AIHW) AIoHaW. Health Expenditure Australia 2006-07. Health and welfare expenditure series no 35 [Internet]. 2007. Available from: http://www.aihw.gov.au/publications/index.cfm/title/10659.

16. Poisal JA. Medicaid drugs. Health Care Financ Rev. 2004;25(3):1-4.
17. Sales MM, Cunningham FE, Glassman PA, Valentino MA, Good CB. Pharmacy benefits management in the Veterans Health Administration: 1995 to 2003. Am J Manag Care. 2005;11(2):104-12.
18. Parliament. House of Commons - Select Committee on Health (Fourth Report) - Control of access to medicines2005. Available from: https://publications.parliament.uk/pa/cm200405/cmselect/cmhealth/42/4209.htm.
19. Sullivan P, Goldmann D. The promise of comparative effectiveness research. JAMA. 2011;305(4):400-1.
20. Smith FJ. Conducting your pharmacy practice research project (second edition): Pharmaceutical Press; 2010.
21. Ades AE, Caldwell DM, Reken S, Welton NJ, Sutton AJ, Dias S. Evidence Synthesis for Decision Making 7. Medical Decision Making. 2013;33(5):679-91.
22. Sackett DL, Strauss SE, Richardson WS, Rosenberg W, Haynes RB. Evidence Based Medicine: How to practice and teach EBM. Edinburgh: Harcourt; 2000.
23. Egger M D-SG, Altman DG. Systematic reviews in health-care: Meta-analysis in context. 2 ed. London: BMJ Publishing Group; 2001.
24. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. PLoS Med. 2007;4(3):e78.
25. Egger M, Juni P, Bartlett C, Holenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. Health Technol Assess. 2003;7(1):1-76.
26. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. Evidence Based Medicine. 2016.
27. Higgins JG, S; . Cochrane Handbook for systematic reviews of interventions. Version 5.0.1. The Cochrane Collaboration, ed2008 2008.
28. The Nordic Cochrance Centre. Review Manager (RevMan) [computer program] [Internet]. 2014. Available from: http://community.cochrane.org/tools/review-production-tools/revman-5/revman-5-download.
29. Suarez-Almazor ME, Belseck E, Homik J, Dorgan M, Ramos-Remus C. Identifying clinical trials in the medical literature with electronic databases: MEDLINE alone is not enough. Control Clin Trials. 2000;21(5):476-87.
30. DeLuca JB, Mullins MM, Lyles CM, Crepaz N, Kay L, Thadiparthi S. Developing a Comprehensive Search Strategy for Evidence Based Systematic Reviews. 2008. 2008;3(1):30.
31. Moher D, Liberati A, Tetzlaff J, Altman DG, Group ftP. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009;339(jul21_1):b2535.
32. Egger M, Zellweger-Zahner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. Lancet. 1997;350(9074):326-9.
33. Baussano I, Brzoska P, Fedeli U, Larouche C, Razum O, Fung IC. Does language matter? A case study of epidemiological and public

health journals, databases and professional education in French, German and Italian. Emerg Themes Epidemiol. 2008;5:16.

34. Morrison A MK, Clark M, Polisena J, Fiander M, Mierzwinski-Urban M, Mensinkai S, Clifford T, Hutton B. English-Language Restriction When Conducting Systematic Review-based Meta-analyses: Systematic Review of Published Studies2009.
35. Dubben H-H, Beck-Bornholdt H-P. Systematic review of publication bias in studies on publication bias. BMJ. 2005;331(7514):433-4.
36. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.
37. Chopra SS. Industry funding of clinical trials: Benefit or bias? JAMA. 2003;290(1):113-4.
38. DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clinical Trials. 1986;7(3):177-88.
39. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Fixed-Effect Versus Random-Effects Models. Introduction to Meta-Analysis: John Wiley & Sons, Ltd; 2009. p. 77-86.
40. Szumilas M. Explaining Odds Ratios. Journal of the Canadian Academy of Child and Adolescent Psychiatry. 2010;19(3):227-9.
41. Kirkwood BR SJ. Medical Statistics. 2 ed: Blackwells Science Ltd; 2003.
42. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. BMJ. 1995;310(6977):452-4.
43. Leung WC. Balancing statistical and clinical significance in evaluating treatment effects. Postgrad Med J. 2001;77(905):201-4.
44. Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011;343.
45. L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. Annals of Internal Medicine. 1987;107(2):224-33.
46. Sterne JA, Egger M, Smith GD. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. BMJ. 2001;323(7304):101-5.
47. Sterne JA, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. J Clin Epidemiol. 2001;54(10):1046-55.
48. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-60.
49. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med. 2004;23(20):3105-24.
50. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ. 2005;331(7521):897-900.
51. Sutton A, Ades AE, Cooper N, Abrams K. Use of indirect and mixed treatment comparisons for technology assessment. Pharmacoeconomics. 2008;26(9):753-67.
52. Coleman CI PO, Cappelleri JC, et al. . Use of Mixed Treatment Comparisons in Systematic Reviews [Internet]2012. Available from: https://www.ncbi.nlm.nih.gov/books/NBK107330/.

53. Lee PM. Bayesian Statistics: An Introduction: Arnold; 1997.
54. Mills EJ, Ghement I, O'Regan C, Thorlund K. Estimating the power of indirect comparisons: a simulation study. PLoS One. 2011;6(1):e16237.
55. Thorlund K, Mills EJ. Sample size and power considerations in network meta-analysis. Syst Rev. 2012;1:41.
56. Mills EJ, Thorlund K, Ioannidis JPA. Demystifying trial networks and network meta-analysis. BMJ : British Medical Journal. 2013;346.
57. Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. Methods in health service research. An introduction to bayesian methods in health technology assessment. BMJ. 1999;319(7208):508-12.
58. Lunn DJ, Thomas A, Best N, Spiegelhalter DJ. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and Computing. 2000;10:325-37.
59. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. Journal of Clinical Epidemiology. 1997;50(6):683-91.
60. Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 2. A Generalised Linear Modelling Framework for Pairwise and Network Meta-analysis of Randomised Controlled Trials. London: NICE; 2011.
61. Dias S SA, Welton NJ. NICE DSU Technical Support Document 3: Heterogeneity: subgroups, meta-regression, bias and bias-adjustment2012 12/06/2014. Available from: http://www.nicedsu.org.uk/TSD3%20Heterogeneity.final%20report.08.05.12.pdf.
62. Dias S. WNJ, Sutton A.J., Ades A.E. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials2014. Available from: http://www.nicedsu.org.uk/TSD2%20General%20meta%20analysis%20corrected%2015April2014.pdf.
63. Jonas DE, Wilkins TM, Bangdiwala S. Findings of Bayesian Mixed Treatment Comparison Meta-Analyses: Comparison and Exploration Using Real-World Trial Data and Simulation [Internet]. . Appendix A, WinBUGS Code Used in Bayesian Mixed Treatment Comparisons Meta-Analysis [Internet]. 2013 12/06/2014. Available from: http://www.ncbi.nlm.nih.gov/books/NBK126104/.
64. Salanti G, Ades AE, Ioannidis JPA. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. Journal of Clinical Epidemiology. 2011;64(2):163-71.
65. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. J Comput Graph Stat. 1998;7:434-55.
66. Lu G, Ades AE. Assessing Evidence Inconsistency in Mixed Treatment Comparisons. Journal of the American Statistical Association. 2006;101(474):447-59.
67. Song F, Clark A, Bachmann MO, Maas J. Simulation evaluation of statistical properties of methods for indirect and mixed treatment comparisons. BMC Medical Research Methodology. 2012;12(1):138.

68. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2002;64(4):583-639.
69. Giavarina D. Understanding Bland Altman analysis. Biochemia Medica. 2015;25(2):141-51.
70. Sedgwick P. Limits of agreement (Bland-Altman method). BMJ : British Medical Journal. 2013;346.
71. Chaimani A, Higgins JPT, Mavridis D, Spyridonos P, Salanti G. Graphical Tools for Network Meta-Analysis in STATA. PLoS ONE. 2013;8(10):e76654.
72. National Institute for Health and Clinical Excellence. Guide to the Methods of Technology Appraisal 2013. London: NICE; 2013.
73. Blume WT. Glossary of Descriptive Terminology for Ictal Semiology: Report of the ILAE Task Force on Classification and Terminology. Epilepsia. 2001;42(9):1212-8.
74. International League Against Epilepsy. Proposal for Revised Classification of Epilepsies and Epileptic Syndromes. Epilepsia. 1989;30(4):389-99.
75. Schmidt D. The clinical impact of new antiepileptic drugs after a decade of use in epilepsy. Epilepsy Res. 2002;50(1-2):21-32.
76. Hauser WA, Hesdorffer DC. Epilepsy: frequency, causes and consequences.: New York: Demos Publications; 1990 1990.
77. Gaitatzis A, Purcell B, Carroll K, Sander JW, Majeed A. Differences in the use of health services among people with and without epilepsy in the United Kingdom: socio-economic and disease-specific determinants. Epilepsy Res. 2002;50(3):233-41.
78. Smith D, Defalla BA, Chadwick DW. The misdiagnosis of epilepsy and the management of refractory epilepsy in a specialist clinic. QJM. 1999;92(1):15-23.
79. Dickersin K, Straus SE, Bero LA. Evidence based medicine: increasing, not dictating, choice. BMJ. 2007;334(suppl_1):s10.
80. Perucca E, Beghi E, Dulac O, Shorvon S, Tomson T. Assessing risk to benefit ratio in antiepileptic drug therapy. Epilepsy Res. 2000;41(2):107-39.
81. Perucca E. Current trends in antiepileptic drug therapy. Epilepsia. 2003;44 Suppl 4:41-7.
82. Mattson RH. Medical management of epilepsy in adults. Neurology. 1998;51(5 Suppl 4):S15-S20.
83. Brodie MJ, Richens A, Yuen AW. Double-blind comparison of lamotrigine and carbamazepine in newly diagnosed epilepsy. UK Lamotrigine/Carbamazepine Monotherapy Trial Group. Lancet. 1995;345:476-9.
84. Brodie MJ, Mumford JP. Double-blind substitution of vigabatrin and valproate in carbamazepine-resistant partial epilepsy. 012 Study group. Epilepsy Res. 1999;34(2-3):199-205.
85. Chadwick D, Marson T. Choosing a first drug treatment for epilepsy after SANAD: randomized controlled trials, systematic reviews, guidelines and treating patients. Epilepsia. 2007;48(7):1259-63.
86. Marson AG, Al-Kharusi AM, Alwaidh M, Appleton R, Baker GA, Chadwick DW, et al. The SANAD study of effectiveness of

carbamazepine, gabapentin, lamotrigine, oxcarbazepine, or topiramate for treatment of partial epilepsy: an unblinded randomised controlled trial. The Lancet. 2007;369(9566):1000-15.

87. Marson AG, Al-Kharusi AM, Alwaidh M, Appleton R, Baker GA, Chadwick DW, et al. The SANAD study of effectiveness of valproate, lamotrigine, or topiramate for generalised and unclassifiable epilepsy: an unblinded randomised controlled trial. The Lancet. 2007;369(9566):1016-26.
88. Pohlmann EB, Van PW, Hallstrom Y, Marson A, Kalviainen R, Marovac J, et al. The Komet study: an open-label, randomized, parallel-group trial comparing the efficacy and safety of levetiracetam with sodium valproate and carbamazepine as monotherapy in subjects with newly-diagnosed epilepsy. Epilepsia. 2008;49 Suppl 7:448-9.
89. Marson ZAK. The New Antiepileptic Drugs: A Systematic Review of Their Efficacy and Tolerability. Epilepsia. 1997;38(8):859-80.
90. National Institute for Clinical Excellence (NICE). The epilepsies: The diagnosis and management of the epilepsies in adults and children in primary care and secondary care 2004 [updated 2004; cited 2010 4/21/2010]. Available from: www.nice.org.uk/cg20.
91. National Institute for Clinical Excellence (NICE). Epilepsies: diagnosis and management 2012 [updated 02/2016; cited 2012]. Available from: https://www.nice.org.uk/guidance/cg137
92. Shorvon S. We live in the age of the clinical guideline. Epilepsia. 2006;47(7):1091-3.
93. Marson AG, Kadir ZA,Chadwick DW. New antiepileptic drugs: a systematic review of their efficacy and tolerability. BMJ. 1996;313(7066):1169-74.
94. Wilby J, Kainth A, Hawkins N, Epstein D, McIntosh H, McDaid C, et al. Clinical effectiveness, tolerability and cost-effectiveness of newer drugs for epilepsy in adults: a systematic review and economic evaluation. Health Technol Assess. 2005;9(15):1-157, iii-iv.
95. Azar NJ, bou-Khalil BW. Considerations in the choice of an antiepileptic drug in the treatment of epilepsy. Semin Neurol. 2008;28(3):305-16.
96. Walker MC, Sander JW. Difficulties in extrapolating from clinical trial data to clinical practice: the case of antiepileptic drugs. Neurology. 1997;49(2):333-7.
97. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ. 2009;339(jul21_1):b2700.
98. Whitehead AW. Meta-Analysis of Controlled Clinical Trials. Chichester, England: John Wiley & Sons; 2002 2002.
99. Spiegelhalter DJ AK, Myles JP. Bayesian Approaches to Clinical Trials and Health-care Evaluation. Chichester, England: John Wiley & Sons; 2004 2004.
100. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. Statistical Methods in Medical Research. 2001;10(4):277-303.
101. UK Gabapentin Study Group. Gabapentin in partial epilepsy. The Lancet. 1990;335(8698):1114-7.

102. Sivenius J, Kalviainen R, Ylinen A, Riekkinen PJ. Double-Blind Study of Gabapentin in the Treatment of Partial Seizures. Epilepsia. 1991;32(4):539-42.

103. Anhut H AP, Feuerstein TJ, Sauermann W, Saunders M, Schmidt B. Gabapentin (Neurontin) as Add-on Therapy in Patients with Partial Seizures: A Double-Blind, Placebo-Controlled Study. Epilepsia. 1994;35(4):795-801.

104. Yamauchi T, Kaneko S, Yagi K, Sase S. Treatment of partial seizures with gabapentin: double-blind, placebo-controlled, parallel-group study. Psychiatry Clin Neurosci. 2006;60(4):507-15.

105. Ben-Menachem E, Biton V, D J, Abous, Doty, Rudd D. Efficacy and Safety of Oral Lacosamide as Adjunctive Therapy in Adults with Partial-Onset Seizures. Epilepsia. 2007;48(7):1308-17.

106. Halasz P, Kalviainen R, Mazurkiewicz M, Rosenow F, Doty P, Hebert D, Sullivan T. Adjunctive lacosamide for partial-onset seizures: Efficacy and safety results from a randomized controlled trial. Epilepsia. 2009;50(3):443-53.

107. Loiseau P, Yuen, A. W., Duche, B., Menager, T., Arne-Bes, M. C. A randomised double-blind placebo-controlled crossover add-on trial of lamotrigine in patients with treatment-resistant partial seizures. Epilepsy Res. 1990;7(2):136-45.

108. Messenheimer J, Ramsay RE, Willmore LJ, Leroy RF, Zielinski JJ, Mattson R, Pellock JM, Valakas AM, Womble G, Risner M. Lamotrigine therapy for partial seizures: a multicenter, placebo-controlled, double-blind, cross-over trial. Epilepsia. 1994;35(1):113-21.

109. Matsuo F, Bergen D, Faught E, Messenheimer JA, Dren AT, Rudd GD, Lineberry CG. Placebo-controlled study of the efficacy and safety of lamotrigine in patients with partial seizures. U.S. Lamotrigine Protocol 0.5 Clinical Trial Group. Neurology. 1993;43(11):2284-91.

110. Biton V, Sackellares JC, Vuong A, Hammer AE, Barrett PS, Messenheimer JA. Double-blind, placebo-controlled study of lamotrigine in primary generalized tonic-clonic seizures. Neurology. 2005;65:1737-43.

111. Naritoku DK, Warnock CR, Messenheimer JA, Borgohain R, Evers S, Guekht AB, et al. Lamotrigine extended-release as adjunctive therapy for partial seizures. Neurology. 2007;69(16):1610-8.

112. Baulac M, Leon, T., O'Brien, T. J., Whalen, E., Barrett, J. A comparison of pregabalin, lamotrigine, and placebo as adjunctive therapy in patients with refractory partial-onset seizures. Epilepsy Res. 2010;91(1):10-9.

113. Cereghino JJ, Biton V, bou-Khalil B, Dreifuss F, Gauer LJ, Leppik I. Levetiracetam for partial seizures: results of a double-blind, randomized clinical trial. Neurology. 2000;55(2):236-42.

114. Shorvon SD, Lowenthal A, Janz D, Bielen E, Loiseau P. Multicenter Double-Blind, Randomized, Placebo-Controlled Trial of Levetiracetam as Add-On Therapy in Patients with Refractory Partial Seizures. Epilepsia. 2000;41(9):1179-86.

115. Betts T, Waegemans T, Crawford P. A multicentre, double-blind, randomized, parallel group study to evaluate the tolerability and efficacy of two oral doses of levetiracetam, 2000 mg daily and 4000 mg daily, without titration in patients with refractory epilepsy. Seizure. 2000;9(2):80-7.

116. Ben-Menachem E, Falter U. Efficacy and tolerability of levetiracetam 3000 mg/d in patients with refractory partial seizures: a multicenter, double-blind, responder-selected study evaluating monotherapy. European Levetiracetam Study Group. Epilepsia. 2000;41:1276-83.
117. Boon P, Chauvel P, Pohlmann EB, Otoul C, Wroe S. Dose-response effect of levetiracetam 1000 and 2000 mg/day in partial epilepsy. Epilepsy research. 2002;48:77-89.
118. Tsai JJ, Yen DJ, Hsih MS, Chen SS, Hiersemenzel R, Edrich P, et al. Efficacy and safety of levetiracetam (up to 2000 mg/day) in Taiwanese patients with refractory partial seizures: a multicenter, randomized, double-blind, placebo-controlled study. Epilepsia. 2006;47(1):72-81.
119. Peltola J CC, Jiménez F, Litovchenko T, Ramaratnam S, Zaslavaskiy L, Lu Z, Sykes DM. Once-daily extended-release levetiracetam as adjunctive treatment of partial-onset seizures in patients with epilepsy: A double-blind, randomized, placebo-controlled trial. Epilepsia. 2009;50(3):406-14.
120. Zhou B, Zhang Q, Tian L, Xiao J, Stefan H, Zhou D. Effects of levetiracetam as an add-on therapy on cognitive function and quality of life in patients with refractory partial seizures. Epilepsy Behav. 2008;12(2):305-10.
121. Xiao Z, Li JM, Wang XF, Xiao F, Xi ZQ, Lv Y, et al. Efficacy and Safety of Levetiracetam (3,000 mg/Day) as an Adjunctive Therapy in Chinese Patients with Refractory Partial Seizures. European Neurology. 2009;61(4):233-9.
122. Barcs G, Walker EB, Elger CE, Scaramelli A, Stefan H, Sturm Y, et al. Oxcarbazepine placebo-controlled, dose-ranging trial in refractory partial epilepsy. Epilepsia. 2000;41(12):1597-607.
123. French JA, Kugler AR, Robbins JL, Knapp LE, Garofalo EA. Dose-response trial of pregabalin adjunctive therapy in patients with partial seizures. Neurology. 2003;60(10):1631-7.
124. Arroyo S AH, Alan R.K., Caroline M.L., Knapp, L.E., Elizabeth, A.G., Messmer, S.L. Pregabalin Add-on Treatment: A Randomized, Double-blind, Placebo-controlled, Dose Response Study in Adults with Partial Seizures. Epilepsia. 2004;45(1):20-7.
125. Beydoun A, Uthman BM, Kugler AR, Greiner MJ, Knapp LE, Garofalo EA. Safety and efficacy of two pregabalin regimens for add-on treatment of partial epilepsy. Neurology. 2005;64(3):475-80.
126. Lee BI, S Y, SB H, MK K, SA L, SK L, et al. Pregabalin add-on therapy using a flexible, optimized dose schedule in refractory partial epilepsies: A double-blind, randomized, placebo-controlled, multicenter trial. Epilepsia. 2009;50(3):464-74.
127. Uthman BM, Rowan AJ, Ahmann PA, Leppik IE, Schachter SC, Sommerville KW, et al. Tiagabine for complex partial seizures: a randomized, add-on, dose-response trial. Arch Neurol. 1998;55(1):56-62.
128. Kalviainen R, Brodie MJ, Duncan J, Chadwick D, Edwards D, Lyby K. A double-blind, placebo-controlled trial of tiagabine given three-times daily as add-on therapy for refractory partial seizures. Epilepsy research. 1998;30(1):31-40.
129. Tassinari CA, Michelucci R, Chauvel P, Chodkiewicz J, Shorvon S, Henriksen O, et al. Double-blind, placebo-controlled trial of topiramate

(600 mg daily) for the treatment of refractory partial epilepsy. Epilepsia. 1996;37(8):763-8.

130. Ben-Menachem E, Henriksen O, Dam M, Mikkelsen M, Schmidt D, S R, et al. Double-Blind, Placebo-Controlled Trial of Topiramate as Add-on Therapy in Patients with Refractory Partial Seizures. Epilepsia. 1996;37(6):539-43.
131. Faught E, Wilder BJ, Ramsay RE, Reife RA, Kramer LD, Pledger GW, et al. Topiramate placebo-controlled dose-ranging trial in refractory partial epilepsy using 200-, 400-, and 600-mg daily dosages. Topiramate YD Study Group. Neurology. 1996;46(6):1684-90.
132. Privitera M, Fincham R, Penry J, Reife R, Kramer L, Pledger G, et al. Topiramate placebo-controlled dose-ranging trial in refractory partial epilepsy using 600-, 800-, and 1,000-mg daily dosages. Topiramate YE Study Group. Neurology. 1996;46(6):1678-83.
133. Sharief M, Viteri C, Ben-Menachem E, Weber M, Reife R, Pledger G, et al. Double-blind, placebo-controlled study of topiramate in patients with refractory partial epilepsy. Epilepsy Res. 1996;25(3):217-24.
134. Yen DJ, Yu HY, Guo YC, Chen C, Yiu CH, Su MS. A double-blind, placebo-controlled study of topiramate in adult patients with refractory partial epilepsy. Epilepsia. 2000;41(9):1162-6.
135. French JA, Mosier M, Walker S, Sommerville K, Sussman N. A double-blind, placebo-controlled study of Vigabatrin three g/day in patients with uncontrolled complex partial seizures. Neurology. 1996;46(1):54-61.
136. Bruni J, Guberman A, Vachon L, Desforges C. Vigabatrin as add-on therapy for adult complex partial seizures: a double-blind, placebo-controlled multicentre study. The Canadian Vigabatrin Study Group. Seizure. 2000;9(3):224-32.
137. Schmidt D, Jacob R, Loiseau P, Deisenhammer E, Klinger D, Despland A, et al. Zonisamide for add-on treatment of refractory partial epilepsy: a European double-blind trial. Epilepsy research. 1993;15(1):67-73.
138. Faught E, Ayala R, Montouris GG, Leppik IE. Randomized controlled trial of zonisamide for the treatment of refractory partial-onset seizures. Neurology. 2001;57(10):1774-9.
139. Brodie MJ. Zonisamide clinical trials: European experience. Seizure. 2004;13 Suppl 1:S66-S70.
140. Sackellares JC, Ramsay RE, Wilder BJ, Browne TR, III, Shellenberger MK. Randomized, controlled clinical trial of zonisamide as adjunctive treatment for refractory partial seizures. Epilepsia. 2004;45(6):610-7.
141. Brodie MJ, Duncan R, Vespignani H, Solyom A, Bitenskyy V, Lucas C. Dose-dependent safety and efficacy of zonisamide: a randomized, double-blind, placebo-controlled study in patients with refractory partial seizures. Epilepsia. 2005;46(1):31-41.
142. Lindberger M, Alenius M, Frisen L, Johannessen SI, Larsson S, Malmgren K, et al. Gabapentin versus vigabatrin as first add-on for patients with partial seizures that failed to respond to monotherapy: a randomized, double-blind, dose titration study. GREAT Study Investigators Group. Gabapentin in Refractory Epilepsy Add-on Treatment. Epilepsia. 2000;41(10):1289-95.

143. Shorvon SD. The Choice of Drugs and Approach to Drug Treatments in Partial Epilepsy. In: Shorvon SD, Perucca E, D F, E D, editors. The Treatment of Epilepsy. Second ed. Oxford, England: Blackwell Science Ltd; 2004. p. 317-33.
144. Otoul C, Arrigo C, van RK, French JA. Meta-analysis and indirect comparisons of levetiracetam with other second-generation antiepileptic drugs in partial epilepsy. Clin Neuropharmacol. 2005;28(2):72-8.
145. Beyenburg S, Stavem, K., Schmidt, D. Placebo-corrected efficacy of modern antiepileptic drugs for refractory epilepsy: systematic review and meta-analysis. Epilepsia. 2010;51(1):7-26.
146. Rheims S, Perucca, E., Cucherat, M., Ryvlin, P. Factors determining response to antiepileptic drugs in randomized controlled trials. A systematic review and meta-analysis. Epilepsia. 2011;52(2):219-33.
147. Devinsky O, Cramer J. Safety and efficacy of standard and new antiepileptic drugs. Neurology. 2000;55(11 Suppl 3):S5-10.
148. Diamond GA, Kaul S. Prior convictions: bayesian approaches to the analysis and interpretation of clinical megatrials. Journal of the American College of Cardiology. 2004;43(11):1929-39.
149. Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. BMJ. 2003;326(7387):472.
150. Psaty BM, Lumley T, Furberg CD, Schellenbaum G, Pahor M, Alderman MH, et al. Health Outcomes Associated With Various Antihypertensive Therapies Used as First-Line Agents: A Network Meta-analysis. JAMA: The Journal of the American Medical Association. 2003;289(19):2534-44.
151. Singh JA, Christensen R, Wells GA, Suarez-Almazor ME, Buchbinder R, Lopez-Olivo MA, et al. A network meta-analysis of randomized controlled trials of biologics for rheumatoid arthritis: a Cochrane overview. Canadian Medical Association Journal. 2009;181(11):787-96.
152. Thijs V, Lemmens R, Fieuws S. Network meta-analysis: simultaneous meta-analysis of common antiplatelet regimens after transient ischaemic attack or stroke. European Heart Journal. 2008;29(9):1086-92.
153. Mills E, Perri D, Cooper C, Nachega J, Wu P, Tleyjeh I, et al. Antifungal treatment for invasive Candida infections: a mixed treatment comparison meta-analysis. Annals of Clinical Microbiology and Antimicrobials. 2009;8(1):23.
154. Huynh T, Perron S, O'Loughlin J, Joseph L, Labrecque M, Tu JV, et al. Comparison of Primary Percutaneous Coronary Intervention and Fibrinolytic Therapy in ST-Segment-Elevation Myocardial Infarction: Bayesian Hierarchical Meta-Analyses of Randomized Controlled Trials and Observational Studies. Circulation. 2009;119(24):3101-9.
155. European Medicines Agency: Committee for Medicinal Products for Human Use (CHMP). Guideline on Clinical Investigation of Medicinal Products in the Treatment of Epileptic Disorders [Draft] 2009 [updated 2009; cited 2010 4/22/2010]. Available from: http://www.ema.europa.eu/pdfs/human/ewp/056698endraft.pdf.

156. Ioannidis JPA. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. Canadian Medical Association Journal. 2009;181(8):488-93.
157. Salanti G, Kavvoura FK, Ioannidis JPA. Exploring the Geometry of Treatment Networks. Annals of Internal Medicine. 2008;148(7):544-53.
158. Ioannidis JPA, Lau J. Completeness of Safety Reporting in Randomized Trials: An Evaluation of 7 Medical Areas. JAMA: The Journal of the American Medical Association. 2001;285(4):437-43.
159. Moher D, Pham, B., Lawson, M. L., Klassen, T. P. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. Health Technol Assess. 2003;7(41):1-90.
160. French JA, Kanner AM, Bautista J, bou-Khalil B, Browne T, Harden CL, et al. Efficacy and tolerability of the new antiepileptic drugs, I: Treatment of new-onset epilepsy: report of the TTA and QSS Subcommittees of the American Academy of Neurology and the American Epilepsy Society. Epilepsia. 2004;45(5):401-9.
161. Zarin W, Veroniki AA, Nincic V, Vafaei A, Reynen E, Motiwala SS, et al. Characteristics and knowledge synthesis approach for 456 network meta-analyses: a scoping review. BMC Medicine. 2017;15:3.
162. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. Lancet. 2006;367(9524):1747-57.
163. Mansia G, De BG, Dominiczak A, Cifkova R, Fagard R, Germano G, et al. 2007 ESH-ESC Guidelines for the management of arterial hypertension: the task force for the management of arterial hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). Blood Press. 2007;16(3):135-232.
164. Jessup M, Brozena S. Heart Failure. The New England Journal of Medicine. 2003;348(20):2007-18.
165. Timmermans PB, Wong PC, Chiu AT, Herblin WF, Benfield P, Carini DJ, et al. Angiotensin II receptors and angiotensin II receptor antagonists. Pharmacological Reviews. 1993;45(2):205-51.
166. Gradman AH, Arcuri KE, Goldberg AI, Ikeda LS, Nelson EB, Snavely DB, et al. A Randomized, Placebo-Controlled, Double-Blind, Parallel Study of Various Doses of Losartan Potassium Compared With Enalapril Maleate in Patients With Essential Hypertension. Hypertension. 1995;25(6):1345-50.
167. Mimran A, Ruilope L, Kerwin L, Nys M, Owens D, Kassler-Taub K, et al. A randomised, double-blind comparison of the angiotensin II receptor antagonist, irbesartan, with the full dose range of enalapril for the treatment of mild-to-moderate hypertension. J Hum Hypertens. 1998;12(3):203-8.
168. Lacourcie're Y, Brunner H, Irwin R, Karlberg BE, Ramsay LE, Snavely DB, et al. Effects of modulators of the renin-angiotensin-aldosterone system on cough. Journal of Hypertension. 1994;12(12):1387-93.
169. Vauquelin G, Fierens F, Liefde IV. Long-lasting angiotensin type 1 receptor binding and protection by candesartan: comparison with other biphenyl-tetrazole sartans. Journal of Hypertension. 2006;24.

170. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ. 2009;339:b2700.

171. Ades AE, Lu G, Higgins JPT. The Interpretation of Random-Effects Meta-Analysis in Decision Models. Medical Decision Making. 2005;25(6):646-54.

172. Andersson OK, Neldam S. The antihypertensive effect and tolerability of candesartan cilexetil, a new generation angiotensin II antagonist, in comparison with losartan. Blood Press. 1998;7(1):53-9.

173. Baguet JP, Nisse-Durgeat S, Mouret S, Asmar R, Mallion JM. A placebo-controlled comparison of the efficacy and tolerability of candesartan cilexetil, 8 mg, and losartan, 50 mg, as monotherapy in patients with essential hypertension, using 36-h ambulatory blood pressure monitoring. Int J Clin Pract. 2006;60(4):391-8.

174. Bakris G, Gradman A, Reif M, Wofford M, Munger M, Harris S, et al. Antihypertensive efficacy of candesartan in comparison to losartan: the CLAIM study. J Clin Hypertens (Greenwich). 2001;3(1):16-21.

175. Gradman AH, Lewin A, Bowling BT, Tonkon M, Deedwania PC, Kezer AE, et al. Comparative effects of candesartan cilexetil and losartan in patients with systemic hypertension. Candesartan Versus Losartan Efficacy Comparison (CANDLE) Study Group. Heart Dis. 1999;1(2):52-7.

176. Vidt DG, White WB, Ridley E, Rahman M, Harris S, Vendetti J, et al. A forced titration study of antihypertensive efficacy of candesartan cilexetil in comparison to losartan: CLAIM Study II. Journal of human hypertension. 2001;15:475-80.

177. Lacourciere Y, Asmar R. A comparison of the efficacy and duration of action of candesartan cilexetil and losartan as assessed by clinic and ambulatory blood pressure after a missed dose, in truly hypertensive patients: a placebo-controlled, forced titration study. Candesartan/Losartan study investigators. Am J Hypertens. 1999;12(12 Pt 1-2):1181-7.

178. Manolis AJ, Grossman E, Jelakovic B, Jacovides A, Bernhardi DC, Cabrera WJ, et al. Effects of losartan and candesartan monotherapy and losartan/hydrochlorothiazide combination therapy in patients with mild to moderate hypertension. Losartan Trial Investigators. Clin Ther. 2000;22(10):1186-203.

179. Willemsen JM, Rabelink TJ, Boer P, Gaillard CA. Disparate systemic and renal blocking properties of angiotensin II antagonists during exogenous angiotensin II administration: implications for treatment. J Hum Hypertens. 2004;18(12):857-63.

180. Granger CB, McMurray JJ, Yusuf S, Held P, Michelson EL, Olofsson B, et al. Effects of candesartan in patients with chronic heart failure and reduced left-ventricular systolic function intolerant to angiotensin-converting-enzyme inhibitors: the CHARM-Alternative trial. Lancet. 2003;362(9386):772-6.

181. McMurray JJ, Ostergren J, Swedberg K, Granger CB, Held P, Michelson EL, et al. Effects of candesartan in patients with chronic heart failure and reduced left-ventricular systolic function taking angiotensin-converting-enzyme inhibitors: the CHARM-Added trial. Lancet. 2003;362(9386):767-71.

182. Granger CB, Ertl G, Kuch J, Maggioni AP, McMurray J, Rouleau JL, et al. Randomized trial of candesartan cilexetil in the treatment of patients with congestive heart failure and a history of intolerance to angiotensin-converting enzyme inhibitors. Am Heart J. 2000;139(4):609-17.

183. Konstam MA, Neaton JD, Dickstein K, Drexler H, Komajda M, Martinez FA, et al. Effects of high-dose versus low-dose losartan on clinical outcomes in patients with heart failure (HEAAL study): a randomised, double-blind trial. Lancet. 2009;374(9704):1840-8.

184. Pitt B, Poole-Wilson PA, Segal R, Martinez FA, Dickstein K, Camm AJ, et al. Effect of losartan compared with captopril on mortality in patients with symptomatic heart failure: randomised trial--the Losartan Heart Failure Survival Study ELITE II. Lancet. 2000;355(9215):1582-7.

185. Pitt B, Segal R, Martinez FA, Meurers G, Cowley AJ, Thomas I, et al. Randomised trial of losartan versus captopril in patients over 65 with heart failure (Evaluation of Losartan in the Elderly Study, ELITE). The Lancet. 1997;349(9054):747-52.

186. BMJ and RPS Publishing. British National Formulary (BNF) 2009.

187. TEVA UK Limited. TEVA launches losartan. 2010.

188. National Institute for Health and Clinical Excellence (NICE). Clinical Guideline 34: Hypertension - Management of hypertension in adults in primary care2006 01/09/2010]. Available from: http://guidance.nice.org.uk/CG34.

189. Weir RA, McMurray JJ, Puu M, Solomon SD, Olofsson B, Granger CB, et al. Efficacy and tolerability of adding an angiotensin receptor blocker in patients with heart failure already receiving an angiotensin-converting inhibitor plus aldosterone antagonist, with or without a beta blocker. Findings from the Candesartan in Heart failure: Assessment of Reduction in Mortality and morbidity (CHARM)-Added trial. Eur J Heart Fail. 2008;10(2):157-63.

190. Staessen JA, Richart T, Wang Z, Thijs L. Implications of recently published trials of blood pressure-lowering drugs in hypertensive or high-risk patients. Hypertension. 2010;55(4):819-31. doi: 10.1161/HYPERTENSIONAHA.108.122879. Epub 2010 Mar 8.

191. Reboldi G, Angeli F, Cavallini C, Gentile G, Mancia G, Verdecchia P. Comparison between angiotensin-converting enzyme inhibitors and angiotensin receptor blockers on the risk of myocardial infarction, stroke and death: a meta-analysis. J Hypertens. 2008;26(7):1282-9. doi: 10.097/HJH.0b013e328306ebe2.

192. National Institute for Health and Clinical Excellence (NICE). Clinical Guideline 108 (update): Chronic Heart Failure2010 21/12/2010. Available from: http://www.nice.org.uk/nicemedia/live/13099/50514/50514.pdf.

193. The ONTARGET Investigators. Telmisartan, Ramipril, or Both in Patients at High Risk for Vascular Events. The New England Journal of Medicine. 2008;358(15):1547-59.

194. Shibata MC, Tsuyuki RT, Wiebe N. The effects of angiotensin-receptor blockers on mortality and morbidity in heart failure: a systematic review. International Journal of Clinical Practice. 2008;62(9):1397-402.

195. Lakhdar R, Al-Mallah MH, Lanfear DE. Safety and tolerability of angiotensin-converting enzyme inhibitor versus the combination of angiotensin-converting enzyme inhibitor and angiotensin receptor blocker in patients with left ventricular dysfunction: a systematic review

and meta-analysis of randomized controlled trials. J Card Fail. 2008;14(3):181-8. doi: 10.1016/j.cardfail.2007.11.008.

196. Haylen BT, de Ridder D, Freeman RM, Swift SE, Berghmans B, Lee J, et al. An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for female pelvic floor dysfunction. Neurourol Urodyn. 2010;29(1):4-20.
197. Milsom I, Abrams P, Cardozo L, Roberts RG, Thuroff J, Wein AJ. How widespread are the symptoms of an overactive bladder and how are they managed? A population-based prevalence study. BJU Int. 2001;87(9):760-6.
198. Abrams P, Kelleher CJ, Kerr LA, Rogers RG. Overactive bladder significantly affects quality of life. Am J Manag Care. 2000;6(11 Suppl):S580-90.
199. Klotz T, Bruggenjurgen B, Burkart M, Resch A. The economic costs of overactive bladder in Germany. Eur Urol. 2007;51(6):1654-62; discussion 62-3.
200. Wein AJ, Rackley RR. Overactive bladder: a better understanding of pathophysiology, diagnosis and management. J Urol. 2006;175(3 Pt 2):S5-10.
201. National Institute for Health and Clinical Excellence (NICE). Urinary incontinence: The management of urinary incontinence in women2013 12/06/2014. Available from: https://www.nice.org.uk/guidance/cg171.
202. McNamara ER, Scales CD. Role of systematic reviews and meta-analysis in evidence-based clinical practice. Indian Journal of Urology : IJU : Journal of the Urological Society of India. 2011;27(4):520-4.
203. Chapple CR, Khullar V, Gabriel Z, Muston D, Bitoun CE, Weinstein D. The effects of antimuscarinic treatments in overactive bladder: an update of a systematic review and meta-analysis. Eur Urol. 2008;54(3):543-62.
204. Buser N, Ivic S, Kessler TM, Kessels AG, Bachmann LM. Efficacy and adverse events of antimuscarinics for treating overactive bladder: network meta-analyses. Eur Urol. 2012;62(6):1040-60.
205. Maman K, Aballea S, Nazir J, Desroziers K, Neine ME, Siddiqui E, et al. Comparative efficacy and safety of medical treatments for the management of overactive bladder: a systematic literature review and mixed treatment comparison. Eur Urol. 2014;65(4):755-65.
206. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009;339:b2535.
207. Lunn D, Thomas A, Best N, Spiegelhalter D. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and Computing. 2000;10(4):325-37.
208. Chapple CR, Dvorak V, Radziszewski P, Van Kerrebroeck P, Wyndaele JJ, Bosman B, et al. A phase II dose-ranging study of mirabegron in patients with overactive bladder. Int Urogynecol J. 2013;24(9):1447-58.
209. Khullar V, Amarenco G, Angulo JC, Cambronero J, Hoye K, Milsom I, et al. Efficacy and tolerability of mirabegron, a beta(3)-adrenoceptor agonist, in patients with overactive bladder: results from a randomised European-Australian phase 3 trial. Eur Urol. 2013;63(2):283-95.

210. Nitti VW, Auerbach S, Martin N, Calhoun A, Lee M, Herschorn S. Results of a randomized phase III trial of mirabegron in patients with overactive bladder. J Urol. 2013;189(4):1388-95.

211. Herschorn S, Barkin J, Castro-Diaz D, Frankel JM, Espuna-Pons M, Gousse AE, et al. A phase III, randomized, double-blind, parallel-group, placebo-controlled, multicentre study to assess the efficacy and safety of the beta(3) adrenoceptor agonist, mirabegron, in patients with symptoms of overactive bladder. Urology. 2013;82(2):313-20.

212. Yamaguchi O, Marui E, Kakizaki H, Homma Y, Igawa Y, Takeda M, et al. Phase III, randomised, double-blind, placebo-controlled study of the beta3-adrenoceptor agonist mirabegron, 50 mg once daily, in Japanese patients with overactive bladder. BJU Int. 2014;113(6):951-60.

213. Appell RA, Sand P, Dmochowski R, Anderson R, Zinner N, Lama D, et al. Prospective randomized controlled trial of extended-release oxybutynin chloride and tolterodine tartrate in the treatment of overactive bladder: results of the OBJECT Study. Mayo Clin Proc. 2001;76(4):358-63.

214. Cardozo L, Lisec M, Millard R, van Vierssen Trip O, Kuzmin I, Drogendijk TE, et al. Randomized, double-blind placebo controlled trial of the once daily antimuscarinic agent solifenacin succinate in patients with overactive bladder. J Urol. 2004;172(5 Pt 1):1919-24.

215. Chapple CR, Rechberger T, Al-Shukri S, Meffan P, Everaert K, Huang M, et al. Randomized, double-blind placebo- and tolterodine-controlled trial of the once-daily antimuscarinic agent solifenacin in patients with symptomatic overactive bladder. BJU Int. 2004;93(3):303-10.

216. Chu F, Smith N, Uchida T. Efficacy and safety of solifenacin succinate 10 mg once Daily: A multicenter, phase III, randomized, double-blind, placebo-controlled, parallel-group trial in patients with overactive bladder. Curr Ther Res Clin Exp. 2009;70(6):405-20.

217. Herschorn S, Heesakkers J, Castro-Diaz D, Wang JT, Brodsky M, Guan Z. Effects of tolterodine extended release on patient perception of bladder condition and overactive bladder symptoms*. Curr Med Res Opin. 2008;24(12):3513-21.

218. Herschorn S, Swift S, Guan Z, Carlsson M, Morrow JD, Brodsky M, et al. Comparison of fesoterodine and tolterodine extended release for the treatment of overactive bladder: a head-to-head placebo-controlled trial. BJU Int. 2010;105(1):58-66.

219. Homma Y, Paick JS, Lee JG, Kawabe K. Clinical efficacy and tolerability of extended-release tolterodine and immediate-release oxybutynin in Japanese and Korean patients with an overactive bladder: a randomized, placebo-controlled trial. BJU Int. 2003;92(7):741-7.

220. Kaplan SA, Schneider T, Foote JE, Guan Z, Carlsson M, Gong J. Superior efficacy of fesoterodine over tolterodine extended release with rapid onset: a prospective, head-to-head, placebo-controlled trial. BJU Int. 2011;107(9):1432-40.

221. Nitti VW, Dmochowski R, Sand PK, Forst HT, Haag-Molkenteller C, Massow U, et al. Efficacy, safety and tolerability of fesoterodine for overactive bladder syndrome. J Urol. 2007;178(6):2488-94.

222. Staskin D, Sand P, Zinner N, Dmochowski R. Once daily trospium chloride is effective and well tolerated for the treatment of overactive

bladder: results from a multicenter phase III trial. J Urol. 2007;178(3 Pt 1):978-83; discussion 83-4.

223. Van Kerrebroeck P, Kreder K, Jonas U, Zinner N, Wein A. Tolterodine once-daily: superior efficacy and tolerability in the treatment of the overactive bladder. Urology. 2001;57(3):414-21.

224. Yamaguchi O, Marui E, Kakizaki H, Itoh N, Yokota T, Okada H, et al. Randomized, double-blind, placebo- and propiverine-controlled trial of the once-daily antimuscarinic agent solifenacin in Japanese patients with overactive bladder. BJU Int. 2007;100(3):579-87.

225. Yamaguchi O, Nishizawa O, Takeda M, Yoshida M, Choo M-S, Gu Lee J, et al. Efficacy, Safety and Tolerability of Fesoterodine in Asian Patients with Overactive Bladder. LUTS: Lower Urinary Tract Symptoms. 2011;3(1):43-50.

226. Zinner N, Gittelman M, Harris R, Susset J, Kanelos A, Auerbach S. Trospium chloride improves overactive bladder symptoms: a multicenter phase III trial. J Urol. 2004;171(6 Pt 1):2311-5, quiz 435.

227. Zinner N, Susset J, Gittelman M, Arguinzoniz M, Rekeda L, Haab F. Efficacy, tolerability and safety of darifenacin, an M(3) selective receptor antagonist: an investigation of warning time in patients with OAB. Int J Clin Pract. 2006;60(1):119-26.

228. Choo MS, Lee JZ, Lee JB, Kim YH, Jung HC, Lee KS, et al. Efficacy and safety of solifenacin succinate in Korean patients with overactive bladder: a randomised, prospective, double-blind, multicentre study. Int J Clin Pract. 2008;62(11):1675-83.

229. BMJ and RPS Publishing. British National Formulary (BNF) [Internet]. Pharmaceutical Press. 2014.

230. Madhuvrata P, Cody JD, Ellis G, Herbison GP, Hay-Smith EJ. Which anticholinergic drug for overactive bladder symptoms in adults. Cochrane Database Syst Rev. 2012;1:CD005429.

231. Minassian VA, Ross S, Sumabat O, Lovatsis D, Pascali D, Al-Badr A, et al. Randomized trial of oxybutynin extended versus immediate release for women aged 65 and older with overactive bladder: lessons learned from conducting a trial. J Obstet Gynaecol Can. 2007;29(9):726-32.

232. Barkin J, Corcos J, Radomski S, Jammal MP, Miceli PC, Reiz JL, et al. A randomized, double-blind, parallel-group comparison of controlled- and immediate-release oxybutynin chloride in urge urinary incontinence. Clin Ther. 2004;26(7):1026-36.

233. Versi E, Appell R, Mobley D, Patton W, Saltzstein D. Dry mouth with conventional and controlled-release oxybutynin in urinary incontinence. The Ditropan XL Study Group. Obstet Gynecol. 2000;95(5):718-21.

234. Anderson RU, Mobley D, Blank B, Saltzstein D, Susset J, Brown JS. Once daily controlled versus immediate release oxybutynin chloride for urge urinary incontinence. OROS Oxybutynin Study Group. J Urol. 1999;161(6):1809-12.

235. National Collaborating Centre for Chronic Conditions (NCCfC). Atrial fibrillation: national clinical guideline for management in primary and secondary care2006.

236. Kannel WB, Wolf PA, Benjamin EJ, Levy D. Prevalence, incidence, prognosis, and predisposing conditions for atrial fibrillation: population-based estimates. Am J Cardiol. 1998;82(8A):2N-9N.

237. Hart RGB, O.; McBride, R.; Pearce, L. A.;. Antithrombotic therapy to prevent stroke in patients with atrial fibrillation: a meta-analysis. Ann Intern Med. 1999;131(7):492-501.

238. Lee SS, A. C.; Cowie, M. R.; . UK stroke incidence, mortality and cardiovascular risk management 1999-2008: time-trend analysis from the General Practice Research Database. BMJ Open. 2011;1(2):e000269.

239. Soliman EZ, Safford MM, Muntner P, Khodneva Y, Dawood FZ, Zakai NA, Thacker EL, Judd S, Howard VJ, Howard G, Herrington DM, Cushman M. Atrial fibrillation and the risk of myocardial infarction. JAMA Intern Med. 2014;174(1):107-14.

240. Smolina K, Wright FL, Rayner M, Goldacre MJ. Incidence and 30-day case fatality for acute myocardial infarction in England in 2010: national-linked database study. Eur J Public Health. 2012;22(6):848-53.

241. Connock MS, C.; Fry-Smith, A.; Jowett, S.; Fitzmaurice, D.; Moore, D.; Song, F.; . Clinical effectiveness and cost-effectiveness of different models of managing long-term oral anticoagulation therapy: a systematic review and economic modelling. Health Technol Assess. 2007;11(38):iii-iv, ix-66.

242. Pirmohamed MJ, S.; Meakin, S.; Green, C.; Scott, A. K.; Walley, T. J.; Farrar, K.; Park, B. K.; Breckenridge, A. M.; . Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. BMJ. 2004;329(7456):15-9.

243. Excellence; NIfHaC. Clinical Guideline 36: National cost impact report to accompany 'Atrial fibrillation: the management of atrial fibrillation'2006.

244. Garcia D, Libby E, Crowther MA. The new oral anticoagulants. Blood. 2010;115(1):15-20.

245. Reilly PAL, T.; Haertter, S.; Connolly, S. J.; Yusuf, S.; Eikelboom, J. W.; Ezekowitz, M. D.; Nehmiz, G.; Wang, S.; Wallentin, L.;. The effect of dabigatran plasma concentrations and patient characteristics on the frequency of ischemic stroke and major bleeding in atrial fibrillation patients: the RE-LY Trial (Randomized Evaluation of Long-Term Anticoagulation Therapy). J Am Coll Cardiol. 2014;63(4):321-8.

246. University of York. Systematic Review: CRDs guidance for undertaking reviews in health care. University of York website: Centre for Reviews and Dissemination; 2009. Available from: http://www.york.ac.uk/crd/SysRev/!SSL!/WebHelp/SysRev3.htm.

247. Assiri A, Al-Majzoub O, Kanaan AO, Donovan JL, Silva M. Mixed treatment comparison meta-analysis of aspirin, warfarin, and new anticoagulants for stroke prevention in patients with nonvalvular atrial fibrillation. 2013;35:967-84.e2.

248. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. 2013;33:607-17.

249. The R Foundation. The R Project for Statistical Computing. Available from: www.r-project.org/

250. National Institute for Health and Clinical Excellence (NICE). Clinical Guideline 36. National Cost Impact Report to accompany ‘Atrial Fibrillation: the Management of Atrial Fibrillation’. London: NICE; 2006.

251. National Institute for Helath and Clinical Excellence (NICE). Guide to the methods of technology appraisal 20132013; (1). Available from: www.nice.org.uk/process/pmg9.
252. Briggs A, Sculpher M, Claxton K. Decision Modelling for Health Economic Evaluation. Oxford: Oxford University Press; 2006.
253. Bayer p. Single Technology Appraisal (STA) of Rivaroxaban (Xarelto). Newbury: Bayer plc; 2011.
254. Wells G, Coyle D, Cameron C, Steiner S, Coyle K, Kelly S. Safety, Effectiveness, and Cost-Effectiveness of New Oral Anticoagulants Compared with Warfarin in Preventing Stroke and Other Cardiovascular Events in Patients with Atrial Fibrillation. Ottawa, ON: Canadian Collaborative for Drug Safety, Effectiveness and Network Meta-Analysis; 2012.
255. Kansal AR, Sorensen SV, Gani R, Robinson P, Pan F, Plumb JM, et al. Cost-effectiveness of dabigatran etexilate for the prevention of stroke and systemic embolism in UK patients with atrial fibrillation. 2012;98:573-8.
256. Shah SV, Gage BF. Cost-effectiveness of dabigatran for stroke prophylaxis in atrial fibrillation. 2011;123:2562-70.
257. Wisloff T, Tove R, Gunhild H, Asmund R, Marianne K. Efficacy and Cost-effectiveness of New Oral Anticoagulants Compared to Warfarin for the Prevention of Stroke in Patients with Atrial Fibrillation. Oslo: Kunnskapssenteret; 2013.
258. Hori M, Connolly SJ, Zhu J, Liu LS, Lau CP, Pais P. Dabigatran versus warfarin: effects on ischemic and hemorrhagic strokes and bleeding in Asians and non-Asians with atrial fibrillation. 2013;44:1891-6.
259. Ali A, Bailey C, Abdelhafiz AH. Stroke prophylaxis with warfarin or dabigatran for patients with non-valvular atrial fibrillation-cost analysis. 2012;41:681-4.
260. Office for National Statistics. 2011 Census, Population and Household Estimates for England and Wales. London: ONS; 2012.
261. Office for National Statistics. ONS Consumer Price Inflation Index for Medical Services (DKC3) for 2013/14. London: ONS; 2013.
262. Department of Health. NHS Reference Costs 2012–13. London: DH; 2014.
263. Ioannides-Demos LL, Makarounas-Kirchmann K, Ashton E, Stoelwinder J, McNeil JJ. Cost of myocardial infarction to the Australian community: a prospective, multicentre survey. 2010;30:533-43.
264. Luengo-Fernandez R, Yiin GS, Gray AM, Rothwell PM. Population-based study of acute- and long-term care costs after stroke in patients with AF. 2013;8:308-14.
265. Hellemons BS, Langenberg M, Lodder J, Vermeer F, Schouten HJ, Lemmens T. Primary prevention of arterial thromboembolism in non-rheumatic atrial fibrillation in primary care: randomised controlled trial comparing two intensities of coumarin with aspirin. 1999;319:958-64.
266. Petersen P, Boysen G, Godtfredsen J, Andersen ED, Andersen B. Placebo-controlled, randomised trial of warfarin and aspirin for prevention of thromboembolic complications in chronic atrial fibrillation. The Copenhagen AFASAK study. 1989;1:175-9.

267. Gulløv AL, Koefoed BG, Petersen P, Pedersen TS, Andersen ED, Godtfredsen J, et al. Fixed minidose warfarin and aspirin alone and in combination vs adjusted-dose warfarin for stroke prevention in atrial fibrillation: Second Copenhagen Atrial Fibrillation, Aspirin, and Anticoagulation Study. 1998;158:1513-21.
268. Liu X, Huang H, Yu J, Cao G, Feng L, Xu Q. Warfarin compared with aspirin for older Chinese patients with stable coronary heart diseases and atrial fibrillation complications. 2014;52:454-9.
269. Yamaguchi T. A Dose Response Study of Dabigatran Etexilate (BIBR 1048) in Pharmacodynamics and Safety in Patients with Non-valvular Atrial Fibrillation in Comparison to Warfarin. 2010.
270. Chung N, Jeon HK, Lien LM, Lai WT, Tse HF, Chung WS. Safety of edoxaban, an oral factor Xa inhibitor, in Asian patients with non-valvular atrial fibrillation. 2011;105:535-44.
271. Yamashita T, Koretsune Y, Yasaka M, Inoue H, Kawai Y, Yamaguchi T. Randomized, multicenter, warfarin-controlled phase II study of edoxaban in Japanese patients with non-valvular atrial fibrillation. 2012;76:1840-7.
272. Weitz JI, Connolly SJ, Patel I, Salazar D, Rohatagi S, Mendell J. Randomised, parallel-group, multicentre, multinational phase 2 study comparing edoxaban, an oral factor Xa inhibitor, with warfarin for stroke prevention in patients with atrial fibrillation. 2010;104:633-41.
273. Chen KP, Huang CX, Huang DJ, Cao KJ, Ma CS, Wang FZ, et al. Anticoagulation therapy in Chinese patients with non-valvular atrial fibrillation: a prospective, multi-center, randomized, controlled study. 2012;125:4355-60.
274. Lopes RD, Alexander JH, Al-Khatib SM, Ansell J, Diaz R, Easton JD. Apixaban for reduction in stroke and other Thromboembolic events in atrial fibrillation (ARISTOTLE) trial: design and rationale. 2010;159:331-9.
275. Granger CB, Alexander JH, McMurray JJ, Lopes RD, Hylek EM, Hanna M. Apixaban versus warfarin in patients with atrial fibrillation. 2011;365:981-92.
276. Hohnloser SH, Hijazi Z, Thomas L, Alexander JH, Amerena J, Hanna M. Efficacy of apixaban when compared with warfarin in relation to renal function in patients with atrial fibrillation: insights from the ARISTOTLE trial. 2012;33:2821-30.
277. Easton JD, Lopes RD, Bahit MC, Wojdyla DM, Granger CB, Wallentin L. Apixaban compared with warfarin in patients with atrial fibrillation and previous stroke or transient ischaemic attack: a subgroup analysis of the ARISTOTLE trial. 2012;11:503-11.
278. Flaker GC, Hohnloser S, Wojdyla D, Hylek E, Garcia D, Sullivan R. Apixaban is efficacious and safe in patients with atrial fibrillation using concomitant amiodarone: an analysis from the ARISTOTLE trial. 2013;1:E317.
279. Al-Khatib SM, Thomas L, Wallentin L, Lopes RD, Gersh B, Garcia D. Outcomes of apixaban vs. warfarin by type and duration of atrial fibrillation: results from the ARISTOTLE trial. 2013;34:2464-71.
280. Bahit MC, Lopes RD, Wojdyla DM, Hohnloser SH, Alexander JH, Lewis BS. Apixaban in patients with atrial fibrillation and prior coronary artery disease: insights from the ARISTOTLE trial. 2013;170:215-20.

281. Garcia DA, Wallentin L, Lopes RD, Thomas L, Alexander JH, Hylek EM. Apixaban versus warfarin in patients with atrial fibrillation according to prior warfarin use: results from the Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation trial. 2013;166:549-58.
282. McMurray JJ, Ezekowitz JA, Lewis BS, Gersh BJ, van Diepen S, Amerena J. Left ventricular systolic dysfunction, heart failure, and the risk of stroke and systemic embolism in patients with atrial fibrillation: insights from the ARISTOTLE trial. 2013;6:451-60.
283. Alexander JH, Lopes RD, Thomas L, Alings M, Atar D, Aylward P. Apixaban vs. warfarin with concomitant aspirin in patients with atrial fibrillation: insights from the ARISTOTLE trial. 2014;35:224-32.
284. Hylek EM, Held C, Alexander JH, Lopes RD, De Caterina R, Wojdyla DM. Major bleeding in patients with atrial fibrillation receiving apixaban or warfarin: The ARISTOTLE Trial (Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation): Predictors, Characteristics, and Clinical Outcomes. 2014;63:2141-7.
285. Ogawa S, Shinohara Y, Kanmuri K. Safety and efficacy of the oral direct factor Xa inhibitor apixaban in Japanese patients with non-valvular atrial fibrillation – The ARISTOTLE-J study. 2011;75:1852-9.
286. Mant J, Hobbs FD, Fletcher K, Roalfe A, Fitzmauricie D, Lip GY. Warfarin versus aspirin for stroke prevention in an elderly community population with atrial fibrillation (the Birmingham Atrial Fibrillation Treatment of the Aged Study, BAFTA): a randomised controlled trial. 2007;370:493-503.
287. Hu DY, Zhang HP, Sun YH, Jiang LQ, Antithrombotic Therapy in Atrial Fibrillation Study G. The randomized study of efficiency and safety of antithrombotic therapy in nonvalvular atrial fibrillation: warfarin compared with aspirin. 2006;34:295-8.
288. Ruff CT, Giugliano RP, Antman EM, Crugnale SE, Bocanegra T, Mercuri M. Evaluation of the novel factor Xa inhibitor edoxaban compared with warfarin in patients with atrial fibrillation: design and rationale for the Effective aNticoaGulation with factor xA next GEneration in Atrial Fibrillation-Thrombolysis In Myocardial Infarction study 48 (ENGAGE AF-TIMI 48). 2010;160:635-41.
289. Giugliano RP, Ruff CT, Braunwald E, Murphy SA, Wiviott SD, Halperin JL. Edoxaban versus warfarin in patients with atrial fibrillation. 2013;369:2093-104.
290. Hori M, Matsumoto M, Tanahashi N, Momomura S, Uchiyama S, Goto S. Rivaroxaban vs. warfarin in Japanese patients with atrial fibrillation – the J-ROCKET AF study –. 2012;76:2104-11.
291. Ezekowitz MD, Reilly PA, Nehmiz G, Simmers TA, Nagarakanti R, Parcham-Azad K. Dabigatran with or without concomitant aspirin compared with warfarin alone in patients with nonvalvular atrial fibrillation (PETRO Study). 2007;100:1419-26.
292. Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A. Dabigatran versus warfarin in patients with atrial fibrillation. 2009;361:1139-51.
293. Connolly SJ, Ezekowitz MD, Yusuf S, Reilly PA, Wallentin L, Randomized Evaluation of Long-Term Anticoagulation Therapy I. Newly identified events in the RE-LY trial. 2010;363:1875-6.

294. Rocket AF, Study I. Rivaroxaban-once daily, oral, direct factor Xa inhibition compared with vitamin K antagonism for prevention of stroke and Embolism Trial in Atrial Fibrillation: Rationale and Design of the ROCKET AF study. 2010;159:340-7.e1.

295. Patel MR, Mahaffey KW, Garg J, Pan G, Singer DE, Hacke W. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. 2011;365:883-91.

296. Hankey GJ, Patel MR, Stevens SR, Becker RC, Breithardt G, Carolei A. Rivaroxaban compared with warfarin in patients with atrial fibrillation and previous stroke or transient ischaemic attack: a subgroup analysis of ROCKET AF. 2012;11:315-22.

297. Mahaffey KW, Wojdyla D, Hankey GJ, White HD, Nessel CC, Piccini JP. Clinical outcomes with rivaroxaban in patients transitioned from vitamin K antagonist therapy: a subgroup analysis of a randomized trial. 2013;158(12):861-8.

298. Stroke Prevention in Atrial Fibrillation I. Warfarin versus aspirin for prevention of thromboembolism in atrial fibrillation: Stroke Prevention in Atrial Fibrillation II Study. 1994;343:687-91.

299. Rash A, Downes T, Portner R, Yeo WW, Morgan N, Channer KS. A randomised controlled trial of warfarin versus aspirin for stroke prevention in octogenarians with atrial fibrillation (WASPO). 2007;36:151-6.

300. Connolly S, Pogue J, Hart R, Pfeffer M, Hohnloser S, Chrolavicius S. Clopidogrel plus aspirin versus oral anticoagulation for atrial fibrillation in the Atrial fibrillation Clopidogrel Trial with Irbesartan for prevention of Vascular Events (ACTIVE W): a randomised controlled trial. 2006;367:1903-12.

301. Eikelboom JW, O'Donnell M, Yusuf S, Diaz R, Flaker G, Hart R. Rationale and design of AVERROES: apixaban versus acetylsalicylic acid to prevent stroke in atrial fibrillation patients who have failed or are unsuitable for vitamin K antagonist treatment. 2010;159:348-53.e1.

302. Hohnloser S, Yusuf S, Eikelboom J, Steg G, Atar D, Budaj A. Apixaban in patients with atrial fibrillation and their risk for cardiovascular hospitalization: insights from the AVERROES trial. 2011;32:671.

303. Connolly SJ, Eikelboom J, Joyner C, Diener HC, Hart R, Golitsyn S. Apixaban in patients with atrial fibrillation. 2011;364:806-17.

304. National Institute for H, Clinical E. Atrial Fibrillation: The Management of Atrial Fibrillation. London: NICE; 2006.

305. Lee S, Anglade MW, Pham D, Pisacane R, Kluger J, Coleman CI. Cost-effectiveness of rivaroxaban compared to warfarin for stroke prevention in atrial fibrillation. 2012;110:845-51.

306. Kamel H, Johnston SC, Easton JD, Kim AS. Cost-effectiveness of dabigatran compared with warfarin for stroke prevention in patients with atrial fibrillation and prior stroke or transient ischemic attack. 2012;43:881-3.

307. Vanassche T, Lauw MN, Eikelboom JW, Healey JS, Hart RG, Alings M. Risk of ischaemic stroke according to pattern of atrial fibrillation: analysis of 6563 aspirin-treated patients in ACTIVE-A and AVERROES. 2015;36:281-27a.

308. Hohnloser SH, Pajitnev D, Pogue J, Healey JS, Pfeffer MA, Yusuf S, et al. Incidence of stroke in paroxysmal versus sustained atrial fibrillation

in patients taking oral anticoagulation or combined antiplatelet therapy: an ACTIVE W Substudy. 2007;50:2156-61.

309. Harrington AR, Armstrong EP, Nolan PE, Malone DC. Cost-effectiveness of apixaban, dabigatran, rivaroxaban, and warfarin for stroke prevention in atrial fibrillation. 2013;44:1676-81.
310. Lee S, Mullin R, Blazawski J, Coleman CI. Cost-effectiveness of apixaban compared with warfarin for stroke prevention in atrial fibrillation. 2012;7:e47473.
311. Liew A, Eikelboom JW, O'Donnell M, Hart RG. Assessment of anticoagulation intensity and management of bleeding with old and new oral anticoagulants. 2013;29(Suppl. 7):34-44.
312. Yang E. A clinician's perspective: novel oral anticoagulants to reduce the risk of stroke in nonvalvular atrial fibrillation – full speed ahead or proceed with caution? 2014;10:507-22.
313. Fenger-Eriksen C, Münster AM, Grove EL. New oral anticoagulants: clinical indications, monitoring and treatment of acute bleeding complications. 2014;58:651-9.
314. UK New Drugs Online Database. New Drugs Online Report for Ciraparantag.
315. Lu G, DeGuzman FR, Hollenbach SJ, Karbarz MJ, Abe K, Lee G. A specific antidote for reversal of anticoagulation by direct and indirect inhibitors of coagulation factor Xa. 2013;19:446-51.
316. UK New Drugs Online Database. New Drugs Online Report for Andexanet Alfa.
317. Schiele F, van Ryn J, Canada K, Newsome C, Sepulveda E, Park J. A specific antidote for dabigatran: functional and structural characterization. 2013;121:3554-62.
318. UK New Drugs Online Database. New Drugs Online Report for Idarucizumab.
319. Halvorsen S, Ghanima W, Fride Tvete I, Hoxmark C, Falck P, Solli O, et al. A nationwide registry study to compare bleeding rates in patients with atrial fibrillation being prescribed oral anticoagulants. European Heart Journal — Cardiovascular Pharmacotherapy. 2017;3(1):28-36.
320. Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann Intern Med. 2004;141(10):781-8.
321. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ. 2010;340.
322. Stewart LA, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data. Cochrane Working Group. Stat Med. 1995;14(19):2057-79.
323. Furukawa TA, Salanti G, Atkinson LZ, Leucht S, Ruhe HG, Turner EH, et al. Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. BMJ Open. 2016;6(7).
324. Zhou X, Cipriani A, Zhang Y, Cuijpers P, Hetrick SE, Weisz JR, et al. Comparative efficacy and acceptability of antidepressants, psychological interventions, and their combination for depressive disorder in children and adolescents: protocol for a network meta-analysis. BMJ Open. 2017;7(8).

325. Trinquart L, Attiche N, Bafeta A, Porcher R, Ravaud P. Uncertainty in treatment rankings: Reanalysis of network meta-analyses of randomized trials. Annals of Internal Medicine. 2016;164(10):666-73.
326. Anderson KM, Wilson PW, Odell PM, Kannel WB. An updated coronary risk profile. A statement for health professionals. Circulation. 1991;83(1):356-62.
327. Wolf PA, D'Agostino RB, Belanger AJ, Kannel WB. Probability of stroke: a risk profile from the Framingham Study. Stroke. 1991;22(3):312-8.
328. Tzoulaki I, Liberopoulos G, Ioannidis JP. Assessment of claims of improved prediction beyond the Framingham risk score. JAMA. 2009;302(21):2345-52.
329. Cupples L, D'Agostino RB. Framingham Heart Study, 30 year follow-up. The Framingham Study: An Epidemiogical Investigation of Cardiovascular Disease: Washington DC; 1988. p. 65-209.
330. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Med Decis Making. 1993;13(4):322-38.
331. Curtis L. Unit Cost of Health and Social Care. Canterburt, Kent: Personal Social Services Research, University of Kent; 2010.

[BLANK PAGE]

10.0 Appendices

10.1 Appendix I: PRISMA templates

PRISMA 2009 Flow Diagram

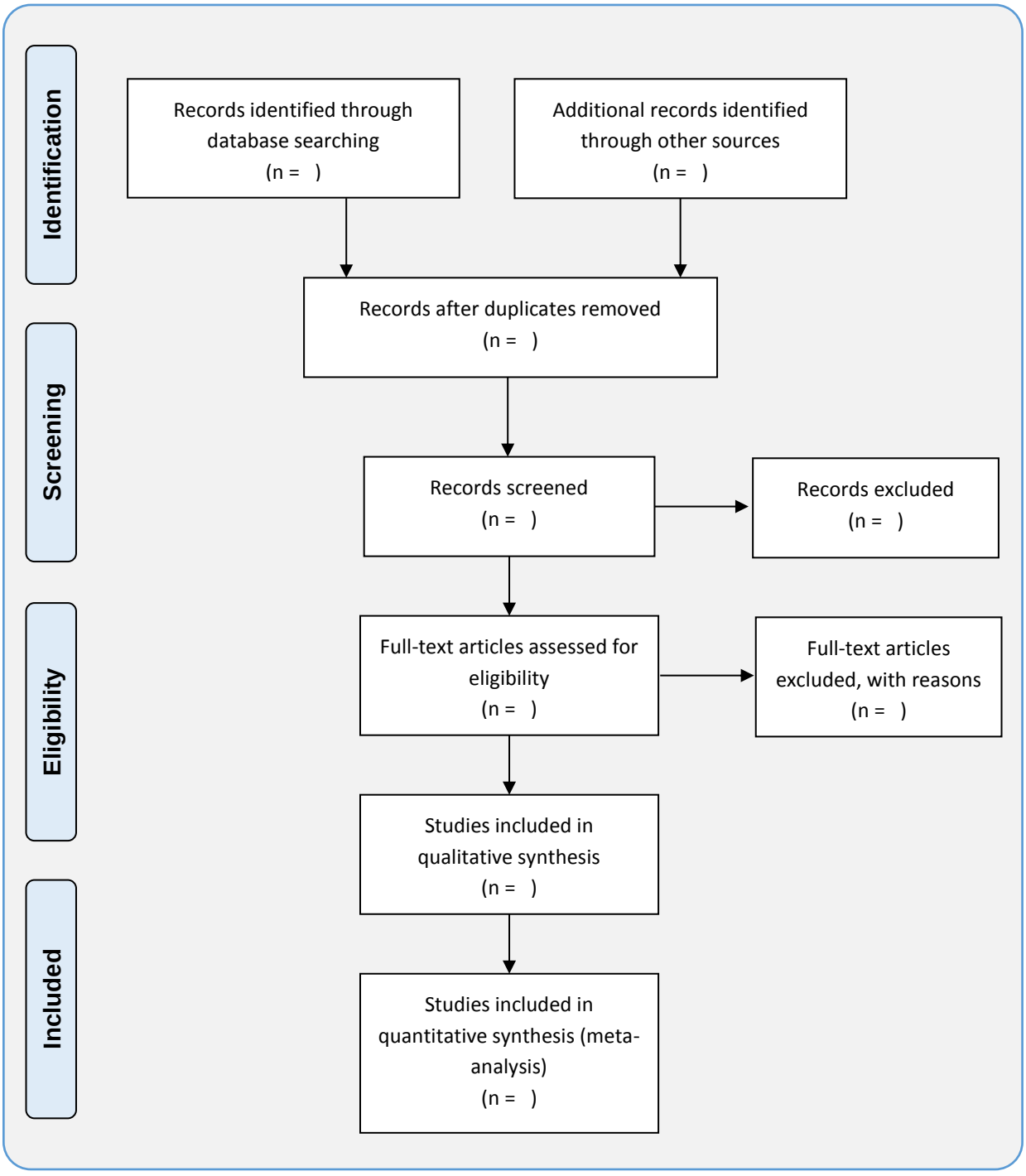


Figure 77: PRISMA flow diagram template (2009)

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	

Section/topic	#	Checklist item	Reported on page #
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	

Section/topic	#	Checklist item	Reported on page #
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	

Figure 78: PRISMA checklist template (2009)

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

10.2 Appendix II: WinBUGS user guide

Instructions for using WinBUGS (v14.0)

Open the programme "WinBUGS14"

Starting the process

[Menu] File – Open – (Open "WinBUGS Template" file)

- A template file with all the commands to run the analysis will open in a window.

Adapt the command section

Replace the following items between the first two blue arrows:

- N (delta) = total number of trial arms (remember, active vs placebo will be entered as two arms) [2]
- NS (mu) = total number of studies [1]
- NT (d) = total number of different treatment option (including placebo which should be number 1) [5]
- BR = baseline risk (i.e. background risk from population data OR pooled result from the placebo intervention) [2]

The program will not run if all of these letters are not replaced with a number (N.B. the BR is in a small font)

Adapt the data

1. Open data sheet in Excel
2. Highlight data only from excel sheet i.e. exclude the trial id, comparison, population (only r[], n[], t[], b[], s[] and m[] and data).
3. Copy
4. Go back to WinBUGS
5. After the second arrow is the data highlight the data (not the word END)
6. [Menu] Edit – Paste Special – Plain Text

Adapt the run program

Below the third arrow is the section "#initial 1"

- Following "d=c(NA,-0.5,-1,-0.2)," repeat the numbers to equal the number of treatments, NA appears once only; e.g. if NT=8 then d=c(NA,-0.5,-1,-0.2,-0.5,-1,-0.2,-0.5)," a comma should be present after the closing bracket.

- Following “mu=c(…)” repeat the numbers to equal the number of studies. Ensure that a comma is not present after the final number before closing the bracket, but is present after the closing bracket.
- Following “delta=c(…)” repeat the numbers to equal the number of trial arms. Ensure that a comma is not present after the final number before closing the bracket. In this entry, a comma is not present after the closing bracket.

Running the program 1

[Menu] Model – Specification

Specification Tool

Step 1

- Highlight the word “model{“ after the first arrow
- On the Specification Tool, Click “check model”
- The status bar (bottom grey bar) should read “model is syntactically correct”

Step 2

- Highlight the row “r[], n[], t[], b[], s[] and m[]”
- On the Specification Tool, click “load data”
- The status bar (bottom grey bar) should read “data loaded”

Step 3

- On the Specification Tool, click “compile”
- The status bar (bottom grey bar) should read “data compiled”

Step 4

- Highlight the row “list(” under #initial 1
- On the Specification Tool, click load inits
- The status bar (bottom grey bar) should read “model is initialized”

Running the program 2

[Menu] Model – Update

Update Tool

Step 5

- Change number of updates from 1000 to 20,000 (number of iterations performed to derive an average.
- Click on “update”
- The iteration box will scan from 0 to 20,000

Running the program 3

[Menu] inference – Sample

Sample Monitor Tool

Step 6

- In the node box, enter the following (not the bit in the brackets), each type pressing set afterwards:
  - or (odds ratio)
  - rr (risk ratio)
  - sumdev (goodness of fit)
  - best (ranking)
  - rk (ranking)
- Go back to the Update Tool
- Click “update”
- The iteration box will scan from 20,000 to 40,000

Executing the program

Go back to the Sample Monitor Tool

Step 7

- In the node box, enter an asterisk (shift+8)
- Click “stats”
- A node statistics window should open will all the results

Obtaining the results

Go to the node statistics window

Step 8

- Place the cursor at the start of the word “node”
- Select all text within the window by pressing “CTRL+A”
- Copy all of the text by pressing “CTRL+C”

Open up a blank Excel spreadsheet

Step 9

- Paste all of the text by pressing “CTRL+V”
- This will copy all the data in columns B to J
- In column A, next to the OR results, enter in the corresponding treatments (for 1, 2, 3, etc.)
- In column K, row 1, add the heading “MCe/sd”
- In K2, setup the rule to divide E2 by D2.
- Drag this rule to apply to the whole of column K
- Using the mouse, left-click on the “K” column button to highlight this column
- Right-click mouse to access the menu and select “format cells”
- Under “number – category” select percentage, changing the number of decimal places to zero before selecting OK.
- The percentages displayed provide a figure (%) on how robust the computed figures are. All values should be less than 5%. If above 5%, go back to WinBUGS, locate the Update Tool, click on update to repeat the number of interations from 40,000 to 60,000. Proceed from Step 7 again.
- Any entries where the sd value is zero should be manually entered as “0%”.

Interpreting the results 1

Goodness of fit (Sumdev)

"An analysis of how the observed value relates to that which is expected"

Step 10

- At the bottom of the data pasted will be a sumdev row, where the number in the C column will indicate how "good" the data is. This number should be in the region of N, the number of studies included in the analysis.
- Below this box (should be empty), enter in the N value
- To calculate a p-value to be more precise in the reporting of this goodness of fit analysis
    - Click on the empty box below the N value (still in column B)
    - Click on the "fx" button in the command bar
    - In the "search for a function box" type in "chi"
    - Select the option "CHIDIST" which will return the one-tailed probability of the chi-squared distribution
    - For the X value, select the sumdev value
    - For the Deg_freedom value, select the N value entered
    - Click OK
    - The p-value should be >0.05 to indicate that there is no statistically significant variation in the result that was observed from that which is expected.

Interpreting the results 2

Switch back to WinBUGS; the node statistics window should still be open.

Click on the Sample Monitor Tool.

Step 11

- In the node box, enter "or"
- Click "history"
- Export these results for analysis; the line should be heavily distorted (up and down) and not a steady line. Copy and paste as appropriate into MS Word.

Step 12

- In the Sample Monitor Tool node box, "or" should still be present
- Click "density"

- A series of graphs, akin to a normal distribution should appear. Copy and paste as appropriate into MS Word.

Interpreting the results 3

Go back to the node statistics window.

[Menu] Inference – Compare

Step 13

- In the node box, enter “or”
- Click “caterpillar”
- A caterpillar plot will be generated representing the odds ratio of all interventions relative to each other visually.
- The red line represents the baseline risk value entered in at the start of the program indicating where interventions need to be placed (with the confidence intervals) to be considered effective at a significant level.

Step 14

- In the node box, enter “rr”
- Click “caterpillar”
- A caterpillar plot will be generated representing the risk ratio of all interventions relative to each other visually.
- Again, the red line represents the baseline risk value entered in at the start of the program indicating where interventions need to be placed (with the confidence intervals) to be considered effective at a significant level. These point estimates take into account all interventions included within the analysis and are the basis for the rank outcome.

Interpreting the results 4

Rank

[Menu] Inference – Rank

- The Rank Monitor Tool window will open
- In the node box, enter “rk”
- Click “set”
- Go back to the Update Tool window and click “update”
- The iteration box will scan from 40,000 to 60,000 (unless repeated earlier in Step 9)
- Go back to the Rank Monitor Tool window
- In the node box, “rk” should still be present
- Click “histogram”

- Information on how the rank was established will be presented using histograms.

Buchers test for inconsistency

Use the worksheet in Excel.

10.2.1 Data entry for 3-arm MTC analysis (Research Project Three)

Table 79: Efficacy endpoint 1; Mean number of micturition's (voids) per 24 hours [change from baseline]

Reference ID	Comparison	t[,1]	t[,2]	t[,3]	t[,4]	y[,1]	y[,2]	y[,3]	y[,4]	se[,1]	se[,2]	se[,3]	se[,4]	na[]
Chapple 2013 (Dragon)	P vs MIR50 v TOL4	1	2	3	NA	-1.44	-2.08	-1.99	NA	0.251	0.191	0.329	NA	3
Khullar 2013 (Scorpio)	P vs MIR50 v TOL4	1	2	3	NA	-1.37	-1.94	-1.57	NA	0.115	0.116	0.123	NA	3
Nitti 2013 (Aries)	P vs MIR50	1	2	NA	NA	-1.05	-1.66	NA	NA	0.130	0.130	NA	NA	2
Herschorn 2013 (Capricorn)	P vs MIR50	1	2	NA	NA	-1.18	-1.6	NA	NA	0.170	0.120	NA	NA	2
Yamaguchi 2014(b) BJUI	P vs MIR50 v TOL4	1	2	3	NA	-0.86	-1.67	-1.4	NA	0.143	0.138	0.134	NA	3
Appell 2001	TOL2 vs OXY10	4	5	NA	NA	-2.87	-3.53	NA	NA	0.003	0.006	NA	NA	2
Cardozo 2004	P vs SOL5 vs SOL10	1	6	7	NA	-1.59	-2.37	-2.81	NA	0.205	0.176	0.199	NA	3
Chapple 2004	P vs SOL5 vs SOL10 vs TOL2	1	6	7	4	-1.2	-2.19	-2.61	-1.88	0.205	0.176	0.199	0.190	4
Choo 2008	SOL5 vs SOL10 vs TOL2	6	7	4	NA	-2.18	-2.47	-2.14	NA	0.249	0.259	0.248	NA	3
Chu 2009	P vs SOL10	1	7	NA	NA	-1.50	-3.00	NA	NA	0.200	0.200	NA	NA	2
Herschorn 2008	P vs TOL4	1	3	NA	NA	-1.70	-2.30	NA	NA	0.200	0.100	NA	NA	2
Herschorn 2010	P vs FES8 vs TOL4	1	8	3	NA	-1.50	-2.20	-2.1	NA	0.200	0.100	0.100	NA	3
Homma 2003	P vs TOL4	1	3	NA	NA	-1.10	-2.00	NA	NA	0.200	0.100	NA	NA	2
Kaplan 2011	P vs FES8 vs TOL4	1	8	3	NA	-2.13	-2.75	-2.48	NA	0.142	0.096	0.107	NA	3
Nitti 2007	P vs FES 4 vs FES8	1	9	8	NA	-1.08	-1.61	-2.09	NA	0.180	0.180	0.180	NA	3
Staskin 2007	P vs TRO60	1	11	NA	NA	-1.99	-2.81	NA	NA	0.160	0.150	NA	NA	2
Van Kerrebroeck 2001	P vs TOL2 vs TOL4	1	4	3	NA	-2.20	-3.30	-3.50	NA	0.178	0.194	0.155	NA	3
Yamaguchi 2007	P vs SOL5 vs SOL10	1	6	7	NA	-0.94	-1.93	-2.19	NA	0.115	0.101	0.109	NA	3
Yamaguchi 2011	P vs FES4 vs FES8	1	9	8	NA	-0.59	-1.15	-1.25	NA	0.180	0.180	0.180	NA	3
Zinner 2004	P vs TRO20	1	10	NA	NA	-1.29	-2.37	NA	NA	0.350	0.350	NA	NA	2
Zinner 2006	P vs DAR15	1	12	NA	NA	-1.80	-2.20	NA	NA	0.350	0.350	NA	NA	2

No studies	21
No treatments	12
No arms	55

Table 81 continued.

Reference ID	Comparison	Events r[]	Sample size n[]	Treatment t[]	Baseline treatment b[]	Study s[]	Arm of study m[]
Chapple 2013 (Dragon)	P vs MIR50 v TOL4	-1.44	166	1	1	1	1
Khullar 2013 (Scorpio)	P vs MIR50 v TOL4	-1.37	480	1	1	2	1
Nitti 2013 (Aries)	P vs MIR50	-1.05	454	1	1	3	1
Herschorn 2013 (Capricorn)	P vs MIR50	-1.18	433	1	1	4	1
Yamaguchi 2014(b) BJUI	P vs MIR50 v TOL4	-0.86	368	1	1	5	1
Appell 2001	TOL2 vs OXY10	-2.87	172	4	4	6	1
Cardozo 2004	P vs SOL5 vs SOL10	-1.59	281	1	1	7	1
Chapple 2004	P vs SOL5 vs SOL10 vs TOL2	-1.20	253	1	1	8	1
Choo 2008	SOL5 vs SOL10 vs TOL2	-2.18	118	6	4	9	1
Chu 2009	P vs SOL10	-1.50	309	1	1	10	1
Herschorn 2008	P vs TOL4	-1.70	201	1	1	11	1
Herschorn 2010	P vs FES8 vs TOL4	-1.50	334	1	1	12	1
Homma 2003	P vs TOL4	-1.10	122	1	1	13	1
Kaplan 2011	P vs FES8 vs TOL4	-2.13	462	1	1	14	1
Nitti 2007	P vs FES 4 vs FES8	-1.08	266	1	1	15	1
Staskin 2007	P vs TRO60	-1.99	300	1	1	16	1
Van Kerrebroeck 2001	P vs TOL2 vs TOL4	-2.20	507	1	1	17	1
Yamaguchi 2007	P vs SOL5 vs SOL10	-0.94	395	1	1	18	1
Yamaguchi 2011	P vs FES4 vs FES8	-0.59	318	1	1	19	1
Zinner 2004	P vs TRO20	-1.29	256	1	1	20	1
Zinner 2006	P vs DAR15	-1.80	225	1	1	21	1

Table 80: Labels and codes used for each of the treatments included within the MTC analysis

Label	Drug (dose)	Code
P	Placebo	1
MIR50	Mirabegron 50mg OD	2
TOL4	Tolterodine MR 4mg OD	3
TOL2	Tolterodine IR 2mg BD	4
OXY10	Oxybutynin MR 10mg OD	5
SOL5	Solifenacin 5mg OD	6
SOL10	Solifenacin 10mg OD	7
FES8	Fesoterodine 8mg OD	8
FES4	Fesoterodine 4mg OD	9
TRO20	Trospium 20mg BD	10
TRO60	Trospium 60mg OD	11
DAR15	Darifenacin 15mg OD	12

```
#Description of data inputs
    # ns = Number of studies
    # nt = Number of treatments (including placebo)
    # t[,x] = Treatment indicator
    # r[,x] = Number achieving response on HAM-D (50% improvement of scores from baseline)
    # n[,x]= Number of all randomized patients (ITT)
    # na[] = Number of arms in study
```

10.3 Appendix III: Supplementary Material for Research Project Two

10.3.1 Cost-Utility Analysis model

Baseline risk parameters (age, gender, SBP), smoking status, presence／absence of diabetes or atrial fibrillation (AF) and high-density lipoprotein (HDL) levels and drug treatment effects on SBP were computed. A risk sub-model was then used to calculate the age- and sex-related probabilities of stroke and CHD risk for each year in the model, based on Framingham risk equations (326, 327). As noted by Wolf et al. (1991), SBP is an accurate predictor of stroke risk (327). Although, other studies have tried to improve the predictive value of outcomes beyond that which the well-established Framingham risk score could predict, a recent review showed that the studies were hampered by methodological flaws (328). Male and female cohorts were modelled separately by baseline risk according to SBP (mild 140 mmHg, moderate 165 mmHg, high 180 mmHg) and subsequent mortality was calculated as a result of myocardial infarction (MI) and cerebrovascular events.

To overcome the normal fixed temporal probability limitation of standard Markov models, annual time-dependent transition probabilities were calculated using look-up tables for age and gender-related all-cause mortality from UK life tables and data from the MI／stroke mortality in the Framingham follow-up Study (329). Mortality was limited to 10 years post disease since the regression equations for stroke progression are valid only for a 10-year period. Half-cycle correction was used when calculating life-years (330).

For a one thousand essential hypertension patient cohort simulation, given varying baseline risks of developing CHD and CVD, we calculated the following secondary economic outcomes: total and average costs per strategy, life-years gained, quality-adjusted life-years (QALYs) gained and report the primary outcome of incremental cost-effectiveness ratios (ICERs) of the two treatments by calculating the ratio of the averaged incremental costs to incremental QALYs. An ICER value of £30,000 per QALY was used as an upper threshold for NHS cost-effectiveness.

Best estimates of disease-state costs and utility values for the base–case model were estimated from the published literature. We considered the direct costs for drug acquisition and those associated with managing initial non-fatal stroke and CHD events as well as the costs for ongoing management, inflated from 2005 to 2009 base year values using the Hospital and Community Pay & Price Index (331). In line with the National Institute for Health and Clinical Excellence (NICE) Technology Appraisals, annual disease-state quality of life was estimated, assuming that MI survivors had higher utility after their initial events and stroke survivors had constant utility thereafter. Cost and health benefits were discounted at a rate of 3.5% in concordance with NICE Technology Appraisal guidance. We assume equal utility loss in both arms attributable to adverse treatment effects since we are comparing two drugs within the same class without proven evidence of any tolerability differences. The death state is associated with zero cost and utility.

10.3.2 UCLH Guidelines on ARBs for Hypertension and Heart Failure

University College London Hospitals NHS
NHS Foundation Trust

UCLH GUIDELINES ON ANGIOTENSIN II RECEPTOR BLOCKERS (ARBs) FOR HYPERTENSION AND HEART FAILURE

In March 2010, losartan (Cozaar®) lost market exclusivity. This is the first of the angiotensin II receptor blockers (ARBs) to do so. This drug class currently incurs the fourth highest cost to the NHS in terms of drug expenditure. The cost of losartan 50mg and 100mg has fallen to £3.77 (from £12.80) and £4.64 (from £16.18) respectively. Experience with other agents (e.g. ACE inhibitors) suggests the price will fall much further saving the NHS more than £50 million. However the total ARB spend in England was £272 million in England in 2008, with 6 ARBs (candesartan, eprosartan, irbesartan, olmesartan, telmisartan and valsartan) remaining on patent. If losartan were the primary ARB in the NHS more than £200 million could be saved per annum. Locally, NHS London spends £35million a year on ARBs generating potential to achieve a £25 million saving.

Within a class of drugs, all acting in the same way at the same receptor, there can be differences in required dose, dosing frequency, drug interactions, adverse event profiles and the pharmacopolitical environment (cost, licensing etc). For ARBs, which are used in hypertension (usually in combination therapy) and heart failure, these differences are small, almost always clinically insignificant, and reflect minor dosing differences. They are all once a day drugs. In other words, the drugs are more or less clinically indistinguishable. With the recent HEAAL trial publication, a target dose for losartan in heart failure is now available at 150mg [unlicensed] per day. After a UCLH Use of Medicines Committee evidence review of ARBs, and in collaboration with Camden PCT, all ARBs except losartan have been removed from The Formulary. For patients on ARBs for hypertension, a pharmacist-led review policy is to be implemented. For patients on ARBs for stable heart failure where the patient is not on the maximum ARB dose, dose escalation is recommended and achieved by changing to losartan. The exception to this is the [rare] specific situation of prior losartan intolerance. For patients with severe heart failure and prone to frequent decompensations, changing medications could theoretically precipitate decompensation (although there is no published evidence that this is as significant problem). However, for such patients the decision to change will be left to the discretion of the patient's physician.

Recommendations

- New patients: Losartan is the only ARB to be initiated in all patients for both hypertension and heart failure. ARBs should only be prescribed if ACE-inhibitor intolerant.
- Existing patients for hypertension:
    - Patients admitted on branded ARBs will receive and be discharged on losartan (pharmacy-led review). Exceptions: prior documented intolerance to losartan.

  Existing patients for heart failure:
    - Where not already on maximal dose, ARB dose escalation is encouraged. If a patient has the ARB dose escalated, the patient

should be changed to losartan (exception: prior documented intolerance to losartan). This should be performed under guidance of the responsible physician. Where the patient is on maximal dose ARB and it is not losartan, the existing ARB is continued and the decision to change left to the responsible physician.

- The losartan dose used in hypertension will be 50mg or 100mg daily depending on whether the existing ARB dose is at the lower or upper end of its dosing schedule (see table). 150mg daily is the target dose for heart failure [unlicensed].

FAQ

Isn't candesartan more effective than losartan in heart failure and /or have a better evidence-base?

There are no head-to-head studies comparing these agents so there is an absence of robust evidence in terms of ascertaining apparent superiority or inferiority of one agent over another. In terms of evidence base, in 2009, a larger study [involving losartan] with longer follow-up, called the HEAAL study, was published (the HEAAL trial recruited 3,846 patients and followed up for a median of almost 5 years; Charm-Alternative [involving candesartan] recruited 2,028 with a median follow-up of about 3 years) which also demonstrated a reduction in the rate of death or admission for heart failure in patients with heart failure, reduced left-ventricular ejection fraction, and intolerance to ACE inhibitors. It is, however, difficult to compare these two trials directly as Charm-Alternative was merely a placebo-controlled trial and only 55% of patients were taking a β -blocker; baseline β-blocker therapy was much greater (72%) in the active-comparator [low dose losartan] HEAAL study. A recent independent analysis concluded that based on these study differences the findings of HEAAL appear reasonably consistent with those of the CHARM-Alternative study.[3]

What about the safety issue raised by some commentators of the ELITE II study, where losartan use was associated with increased mortality compared to placebo in patients prescribed beta-blockers. HEAAL had no placebo arm and cannot therefore rule out the possibility that losartan interacts with beta-blockers in patients with heart failure in a negative fashion?

First, this finding was based on exploratory *post-hoc* sub-group analyses in a small (insufficiently powered) group of patients in ELITE II. Such data are interpreted as hypothesis-generating and not policy-defining. Second, this apparent difference was not seen if use was based on ***concomitant*** treatment with beta-blockers during the study which is more relevant. Third, patients on losartan and captopril also taking beta-blockers did better than most patients not on such treatment at randomisation during ELITE II. Fourth, the "interaction" between treatment effect and ***baseline*** beta-blocker use should be interpreted with caution given the small absolute number of patients receiving these drugs and potential bias related to the reasons for administering these agents. Fifth, the presence of an interaction was investigated in the subsequent OPTIMAAL study (see below) which showed no interaction between losartan and beta blockers with respect to survival. Sixth, many commentators now agree that these "findings" suggested from ELITE II analyses are likely to be spurious. Seventh, of note, baseline beta-blocker therapy was much greater in the HEAAL and OPTIMAAL studies (72 and 79% respectively) when compared to CHARM-Alternative (55%). Finally, there is no evidence from HEAAL of an inflated mortality rate in either arm.

Wasn't losartan inferior to ACEi for post-MI heart failure (OPTIMAAL study)?
5477 patients 50 years of age or older with confirmed acute myocardial infarction and heart failure during the acute phase or a new Q-wave anterior infarction or reinfarction, were randomly assigned and titrated to a [low] target dose of losartan (50 mg once daily) or [proven] captopril (50 mg three times daily) as tolerated in the OPTIMAAL study. The primary endpoint was all-cause mortality. There were 946 deaths during a mean follow-up of 2.7 years: 499 (18%) in the losartan group and 447 (16%) in the captopril group (relative risk 1·13 [95% CI 0·99—1·28], p=0·07) i.e. a non-significant difference in the primary endpoint (total mortality) was observed. The incidence of reinfarction, revascularisation, and all-cause hospital admission were essentially identical between the two groups. The upper one-sided 95% confidence margin (1·25) for the relative risk of death from any cause was above the pre-specified margin of 1·10 resulting in failure to satisfy the non-inferiority criterion.

Valsartan and candesartan lose exclusivity within 2 years, is this worth all the effort?
Patents are often extended and there is often a delay between total expiry and generic availability. Regardless, savings of several hundred million pounds that could be achieved during this time are far from insignificant.

It would be easier to apply these guidelines for new patients only, won't changing patients stabilized on treatment be too time consuming and risky?
For patients treated for hypertension, no experts consulted at UCL Hospitals thought this unreasonable especially considering the wide inter- and intra-day variations in physiological blood pressure; most use will be as combination therapy as well. For patients with heart failure the guideline allows for clinical discretion and only suggests change during up-titration to target dose; a practice that is universally encouraged in heart failure.

Why don't you recommend changing all heart failure patients?
The evidence for all ARBs in heart failure is robust, including evidence for losartan. The effect of ARBs in heart failure is a class effect. There are no head-to-head studies indicating superiority or inferiority of one agent over another. Also, losartan is the only ARB where high dose has been shown to be more effective than low dose. Nevertheless, despite all the pharmacological data of equivalence, there may be small 'dosology' differences between ARBs. Some heart failure patients are very unwell with short life expectancy. It is possible that even slight changes such as may occur during an ARB change could tip patients into decompensation – or, of course, decompensation episodes could occur anyway and be blamed on the ARB change, preventing disclosure of other causes.

High dose losartan for heart failure is not licensed.
This is true, but sometimes the regulatory environment lags the evidence, and sometimes, such as when the drug is off patent, it may never catch up. Doctors use unlicensed drug/dose/indications frequently (such as generic clopidogrel in acute coronary syndrome), provided the reasons are sensible and robust. The UCLH Use of Medicines Committee has also advocated the use of 150mg losartan in heart failure.

Are there any significant differences in drug-drug or drug-food interactions?
No although aliskiren levels may be reduced by irbesartan.

What about combination drugs?
Patients appreciate combination drugs (ARB + diuretic) as it reduces the total number of tablets per day to be taken. However, the price premiums for combination drugs are too high so it is not cost-effective to prescribe them. Any individuals on combination treatment will be changed to losartan plus equivalent dose diuretic.

What are the precise dose comparisons with losartan?
All 7 drugs are marketed in 3 doses (low, medium, high dose). We recommend changing to low, medium or high dose losartan (see table). During changing, there may be the opportunity to uptitrate the dose because higher doses of losartan are better in heart failure.

Limiting medical prescribing is a threat to doctor autonomy.
NHS resources are limited and getting more so. Optimising cost-effective prescribing of ARBs will free up resources for other areas where there may be no alternatives, particularly new, expensive but effective treatments.

What will the saving be?
The savings of an England-wide change are in the region of £200 million per year, and £25million in London. This saving will not, however, be sustained long term as other ARBs lose their patent (e.g. candesartan 2012) – but in an environment of fiscal tightening, this will free up resources for other essential services and drugs over the next few years.

Will patients want to change?
There is extensive experience of changing similar drugs both locally and nationally, e.g. in statins. We have also developed a patient information leaflet to explain the justification and reasons for changing. When conducted sensitively, patients almost universally approve and consent to changing.

Who contributed to these guidelines?
Dr James Moon & Dr Simon Woldman, both Consultant Cardiologists; Dr Anthony Grosso & Mr Pritesh Bodalia, both Pharmacists; and Prof Raymond MacAllister & Prof Aroon Hingorani, both Consultant Clinical Pharmacologists, drafted these guidelines which were approved by the multidisciplinary Use of Medicines Committee (Drugs and Therapeutics Committee).

Appendix

Table 1: Dose comparisons of ARBs in heart failure

ARB	Lower dose	Intermediate dose	Target dose
Losartan	50mg	100mg	150mg [unlicensed]
Candesartan	4mg	8-24mg	32mg
Valsartan	80mg	120-280mg	320mg

Table 2: Dose comparisons of ARBs in hypertension

ARB	Lower dose	Intermediate dose	Upper dose
Losartan	50mg	50-100mg	100mg
Candesartan	4mg	8-24mg	32mg
Eprosartan	600mg	600mg	800mg
Irbesartan	150mg	150mg	300mg
Olmesartan	10mg	20mg	30-40mg
Telmisartan	20mg	40mg	80mg
Valsartan	80mg	120-280mg	320mg

Note: These are not exact dose equivalent tables and lower doses may be appropriate for the elderly, patients with renal or hepatic impairment or if on concomitant diuretic treatment (see current British National Formulary or Summary of Product Characteristics for prescribing information).

References

1. McKelvie, R. Compared with low-dose losartan, high-dose losartan decreases risk of death or hospital admission for heart failure in people with heart failure who are intolerant to ACE inhibitors Evidence-Based Medicine April 2010 | volume 15 | number 2

2. Grosso AM, Bodalia PN, MacAlliser RJ, Hingorani AD, Moon JC, Scott MA. Comparative clinical- and cost-effectiveness of candesartan and losartan in the management of hypertension and heart failure: a systematic review, meta- and cost-utility analysis. Int J Clin Pract 2011; 65 (3): 253-263

10.3.3 GP Template Letter

University College London Hospitals NHS
NHS Foundation Trust

Patient name: _____

DOB: _____

Hospital admission date: _____

Dear Doctor,

Re: Angiotensin II Receptor Blockers (ARBs)

We are writing to inform you that we have changed the above patient to losartan therapy during their recent hospital admission. In March 2010, losartan (Cozaar®) lost its patent. This is the first of the angiotensin II receptor blockers (ARBs) to do so. We estimate that if losartan were the primary ARB in the NHS more than £200 million could be saved per annum. Locally, NHS London spend £35 million a year on ARBs generating the potential to achieve a £25 million saving.

The UCLH Use of Medicines Committee (UMC) which includes representation from our local PCTs, has recently reviewed the evidence for ARBs and it has been agreed that we will change all patients prescribed a branded ARB for hypertension to generic losartan. The exception to this is the [rare] specific situation of prior losartan intolerance. For patients on ARBs for stable heart failure where the patient is not on the maximum ARB dose, we will have changed to losartan during dose escalation (our target dose for heart failure [150mg] is unlicensed, although it is evidence-based). For patients with severe heart failure and prone to frequent decompensations, changing medications could theoretically precipitate decompensation (although there is no published evidence that this is a significant problem). However, for such patients the decision to change has been left to the discretion of the patients' physician.

We have informed the patient of this change and provided them with an information leaflet and 28 days supply of losartan tablets with the instruction to obtain further supplies from your surgery.

If you would like to discuss this further, please contact the Medicines Management team on 0845 1555 000 extension 4340. Patients can phone our Medicines Information helpline on 0845 1555 000 extension 73500.

Yours sincerely

Mr Pritesh Bodalia
On behalf of the UCLH UMC

10.3.4 Heart Failure Patient Information Leaflet (PIL)

University College London Hospitals NHS
NHS Foundation Trust

Patient Information: Changes to your heart failure medication

This leaflet aims to explain the reasons why one of the medications to help with control of your heart failure has been changed. The UCLH Use of Medicines Committee is a group of specialist consultants, pharmacists and nurses that has a special interest in medicines. They regularly review the medicines we prescribe to check that we're using the most effective ones. For conditions for which there are a number of similar medicines we also check that we are using the most cost-effective medicine (see below for more information). This group has recently reviewed the potential benefits of each of the medicines used to treat heart failure and it was agreed that losartan should be used in preference to candesartan, eprosartan, irbesartan, olmesartan, telmisartan or valsartan.

We have changed your prescription to losartan 50mg, 100mg or 150mg to be taken once a day to help with your heart failure. If you have any questions about this change please contact our **Medicines Information** department on **0845 1555 000 extension 73500**. This line is open between **10am to 5pm** from **Monday to Friday**. If you have any candesartan, eprosartan, irbesartan, olmesartan, telmisartan or valsartan at home please return them to your local chemist so that they can be disposed of. **Do not take candesartan, eprosartan, irbesartan, olmesartan, telmisartan or valsartan AND losartan at the same time.**

Some reasons for the change:

For many illnesses there are a number of similar medicines. The more expensive one is not necessarily better than less expensive ones. When there is more than one choice it is sensible for the doctor to prescribe the one that is the best value (or most cost-effective).

We do not believe there is anything medically wrong with the medicine you have been using previously. But we do believe that losartan can give you the same results while also giving the NHS better value for money.

If the new medicine does not suit you, your GP can change your prescription back to the medicine you have been taking.

Frequently asked questions:
Why change my medicine if my current one works fine? Your previous medicine was one of a number of very similar medicines. The newer medicine you have been given is currently the most cost-effective of this class and allows us to use these savings to help improve other NHS services.

Is there any risk of side effects with the change? As with all medicines the risk is not zero but it is a small risk. If you do experience a side effect it is likely to be minor and you are no more likely to experience a side-effect with the new medicine than the old one. If you experience any problems tell your doctor who can review your symptoms and if expected to be a side effect to losartan they can change your prescription back to the one you were using previously.

Will I be asked to change this medicine again? We expect this medicine to be the most cost-effective for a few years so it is unlikely that you would be asked to change again, certainly not for at least a few years.

How much money will this save the NHS? A substantial amount (many millions of pounds). Money that is saved will be used to improve other local NHS services.

10.3.5 Hypertension Patient Information Leaflet (PIL)

University College London Hospitals NHS
NHS Foundation Trust

Patient Information: Changes to your blood pressure medication

This leaflet aims to explain the reasons why one of the medications to lower your blood pressure has been changed. The UCLH Use of Medicines Committee is a group of specialist consultants, pharmacists and nurses that has a special interest in medicines. They regularly review the medicines we prescribe to check that we're using the most effective ones. For conditions for which there are a number of similar medicines we also check that we are using the most cost-effective medicine (see below for more information). This group has recently reviewed the potential benefits of each of the medicines used to lower blood pressure and it was agreed that losartan should be used in preference to candesartan, eprosartan, irbesartan, olmesartan, telmisartan or valsartan.

We have changed your prescription to losartan 50mg or 100mg to be taken once a day to lower your blood pressure. If you have any questions about this change please contact our **Medicines Information** department on **0845 1555 000 extension 73500**. This line is open between **10am to 5pm** from **Monday to Friday**. If you have any candesartan, eprosartan, irbesartan, olmesartan, telmisartan or valsartan at home please return them to your local chemist so that they can be disposed of. **Do not take candesartan, eprosartan, irbesartan, olmesartan, telmisartan or valsartan AND losartan at the same time.**

Some reasons for the change:

For many illnesses there are a number of similar medicines. The more expensive one is not necessarily better than less expensive ones. When there is more than one choice it is sensible for the doctor to prescribe the one that is the best value (or most cost-effective).

We do not believe there is anything medically wrong with the medicine you have been using previously. But we do believe that losartan can give you the same results while also giving the NHS better value for money.

If the new medicine does not suit you, your GP can change your prescription back to the medicine you have been taking.

Frequently asked questions:
Why change my medicine if my current one works fine? Your previous medicine was one of a number of very similar medicines. The newer medicine you have been given is currently the most cost-effective of this class and allows us to use these savings to help improve other NHS services.

Is there any risk of side effects with the change? As with all medicines the risk is not zero but it is a small risk. If you do experience a side effect it is likely to be minor and you are no more likely to experience a side-effect with the new medicine than the old one. If you experience any problems tell your doctor who can review your symptoms and if expected to be a side effect to losartan they can change your prescription back to the one you were using previously.

Will I be asked to change this medicine again? We expect this medicine to be the most cost-effective for a few years so it is unlikely that you would be asked to change again, certainly not for at least a few years.

How much money will this save the NHS? A substantial amount (many millions of pounds). Money that is saved will be used to improve other local NHS services.

10.4 Appendix IV: NOAC systematic review: data extraction guide

Using the Access Database

DATA EXTRACTION GUIDE

Getting started:

- Double-click on the Access database file "Pritesh_NOACs_AF.mdb" or right click and select open to open. The first page displays many tabs: - 'Documentation', 'Dataset Information', 'Data Entry', 'Codes & Recoding', 'Utilities', and 'Data Administrator'.
- Click once on 'Data Entry' tab ignoring others. Three options display under Data Entry.
- Double-click on 'NOACs Data Entry'. Two options display under NOACs Data Entry.
- Double-click on 'Start NOACs Data Entry' to open the database. When the database is open, four tabs will be displayed: Paper Retrieval Admin, Screen Titles and Abstracts, Screen Full Papers, and Data Extraction.
- Click once on 'Data Extraction' tab to start data extraction. The first level of data extraction is for the trial-level data. The form to extract information with is the first form that displays when you click on 'Data Extraction' tab. Displayed horizontally at the middle of the form are five tabs for different sections of trial-level data. Clicking on each tab will open a sub-form with which to extract data for that section.
- Start with the first tab on the far right i.e. Trial data extraction administration tab and work towards the left until complete 'Trial-level risk of bias assessment'.
- The 'trial references' tab displays all of the papers that are related to a specific trial. The 'trial population characteristics' tab displays population level characteristics and will only need to be completed if arm level characteristics are not reported.

Arm-level information and outcomes/results

1. After completing trial-level data extraction, open arm-level data form by clicking on the 'Arm-level data' tab at the top left of the trial level information form. When open, there are two 'computer repair' symbols – a large symbol at the top left of the form, and a smaller symbol at the middle aspect of the form towards the right. First, add study intervention arms for them to be displayed in the vertical box on the right. To add an intervention arm, click on the smaller 'computer repair' symbol and then click on 'add a new event'. A new form will open up. Trial ID, Trial name and Trial registration number are already populated. Add the intervention arm number (as in the study), and select an intervention code and the date of the extraction (i.e. the date that you are doing the extraction). To add other intervention arms, click

on ‘Add Event’. When you finish adding all the intervention arms, click on ‘Add Event’ again to return to the Arm-level information page. The added arms now appear in the vertical box to the right. Click on an intervention arm and then click on ‘Study Details’ tab to add intervention arm specific information for each of the arms. When complete for all intervention arms, close the arm-level form by clicking on the tab marked ‘X’ on the top mid-left corner of the form beside the big ‘computer repair’ symbol.

2. To enter outcomes/results, click on ‘Enter Outcomes and Results’ tab at the top left of the trial level information form (below the ‘Enter Arm-level data’). A similar form to the arm-level form (already completed above) will open. To add outcomes, follow the same process for adding the different intervention arms as described above. When all outcomes have been added, the outcomes will now appear on the vertical box to the right (just like the intervention arms are displayed). Click on an outcome and then click on ‘Outcome details and RoB’ tab to add outcome details and complete risk of bias assessment for that outcome. To add results for an outcome, first select the intervention arms in a study. To do this, select the outcome, click on ‘Results’ tab at the top, then click on add at the top of the vertical box on the right side of the result page that opened up. All intervention arms already stored at ‘1’ above will appear in a drop down. Now select the arm number as in the study. Click ‘Add Result’ at the top of the page to repeat the process and add another intervention arm. When all intervention arms have been added they will all be displayed on the vertical box to the right. Select each intervention arm to add the result for each arm. To exit outcomes and results page after completing the fields, click on the ‘X’ tab at the top left of the main form.

Note: When the data field in a form is clicked once, the description of the field will normally appear at the bottom of the screen as it does in Access version 2003. However, this display may depend on the MS Access version that is in use; so this display may not appear in some versions. Below are important data fields and descriptions.

Data field description

1. Trial Identification: This is a unique id automatically assigned to each trial by the computer.

2. Trial registration no.: This is a unique registration number that published trials have. The field is a text field so for any trial with missing registration number, make up a registration number for it but keep a record of such trials.

3. Trial name: This is a name (usually, but not exclusively an abbreviation) that a trial is known with – for example RE-LY. This is a text field so type in the name as it is on publication.

4. Trial country: This is a list of all countries where a trial was conducted. It is a memo field so add all countries.
5. Trial description: This is a description of a trial. This is a text field.
6. Trial dates: This is the commencement & end dates for a trial; example "1970 to 1973".
7. Data extracted by: This is the identification code for a reviewer who is extracting data (id codes are selected from the drop down menu).
8. Extraction date: This is the date that data was extracted for this research project.
9. Data checked by: This is identification code for the reviewer who checked extracted data.
10. Check date: This is the date on which data extraction was independently checked.
11. Data agreed and final: This is whether data extraction discrepancies have been resolved.
12. Was this paper used: This informs whether data was extracted from a paper or not.
13. Number of trial centres: This is the number of centres where a trial was conducted.
14. Trial setting: This is the setting where a trial was conducted. For example, hospital etc.
15. Number screened: This is the number of all screened individuals for consideration in a trial.
16. Number eligible: This is the number of eligible individuals after screening.
17. Number randomised: This is the total number of participants randomised in a trial.
18. List all composite outcomes reported in this trial: This is a list of all composite outcomes reported in a study, if any.
19. Inclusion criteria: This is the specified criteria used to include study participants. This is a text field.
20. Exclusion criteria: This is the specified criteria for exclusion of individuals from a study. This is a text field.
21. Mean age of population, SD mean age, Measure bleed risk, Mean bleed risk, SD mean bleed risk, Mean BMI for population, SD mean BMI, Mean weight for population, SD mean weight, Male %, %

Previous stroke/TIA, % Previous thromboembolic event, % Previous MI, % Hypertension, % Diabetes Mellitus, % Liver disease, % Renal disease, % Pregnancy, % CHF.

(The above are self-descriptive and should be extracted as documented in a trial)

22. Ethnicity information: This is the percentage of different ethnic groups among participants as reported in a trial. For example, %White, %Black, %Hispanic, etc. This is a text field.

23. AF type: This is the type of atrial fibrillation (by %) among the trial population. For example, Paroxysmal, Persistent, Permanent, New onset, Mixed. This is a text field.

24. Index event: This is the index event in the study population. For example, DVT, PE, etc. Complete this for secondary prevention of VTE only.

25. Arm Number: This is the number given to an intervention arm in a trial.

(Please ensure that the correct arm results are documented for an intervention arm)

26. Outcome Type: This is whether an outcome is single outcome or composite.

27. Outcome Definition: This is the definition of an outcome in a trial. You could copy from a trial paper and paste in the field.

28. Treatment duration: This is the duration of intervention administered. It is a text field. (Please specify whether the duration is in days, weeks or months)

29. Time to outcome measure: This is the time from baseline (start of intervention) to when an outcome was measured; in weeks, months, years (please specify). This is a text field.

30. Location of results: This is the location of extracted result in a paper. This is a text field.

Risk of bias assessment

1. Statement: This is the supporting statements for a reviewer's judgment on a specific risk of bias. This is a text field.

2. Quote: This is the supporting quote from a paper for judgment on a specific risk of bias. The quote should be copied from a paper and pasted into the field. This is a text field.

3. Quote location: This is the location in a paper from which a quote is copied. For example, "third paragraph on page 23". This is a text field.

10.5 Appendix V: NOAC systematic review: database screenshots

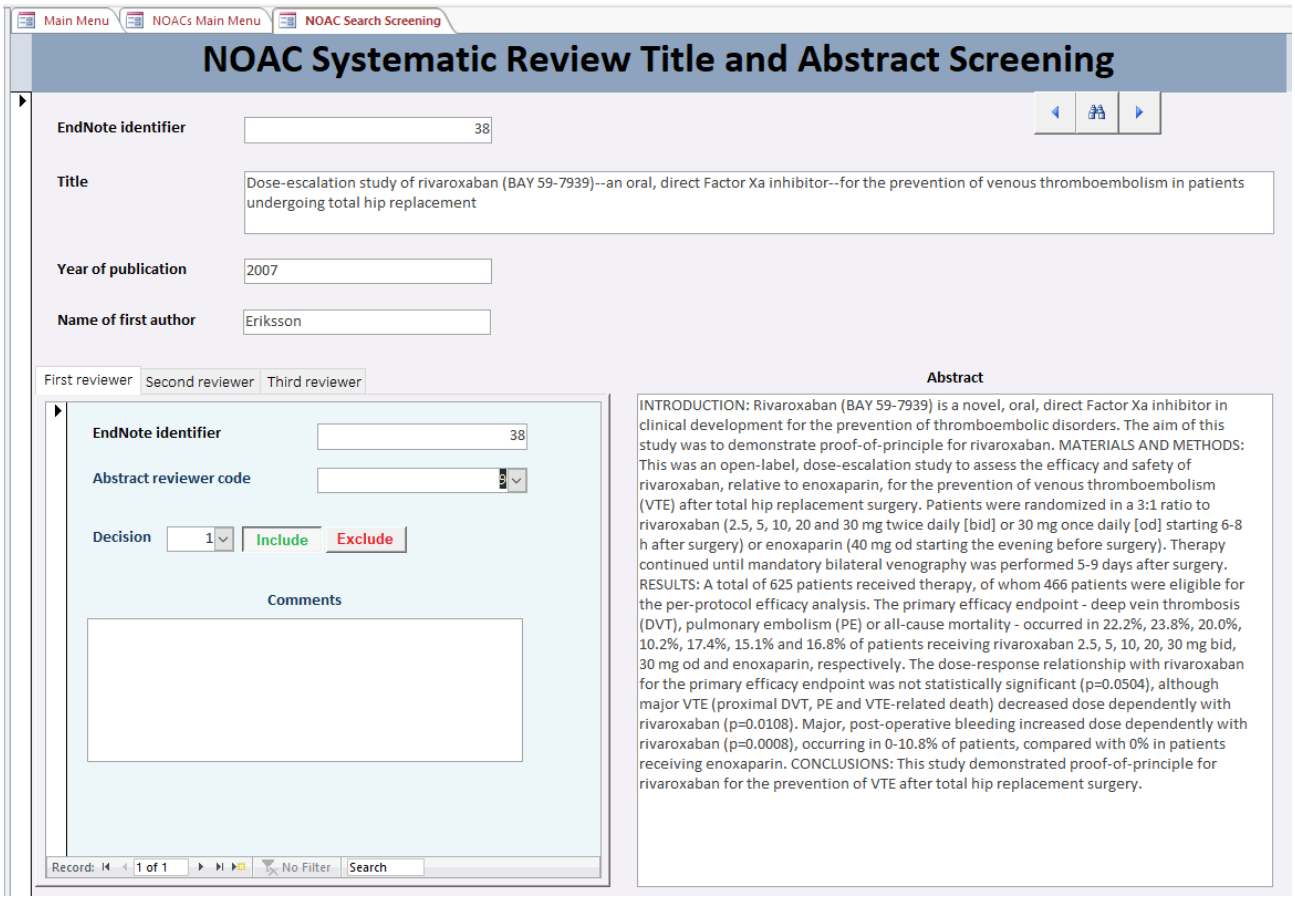


Figure 79: Screening for inclusion at abstract level

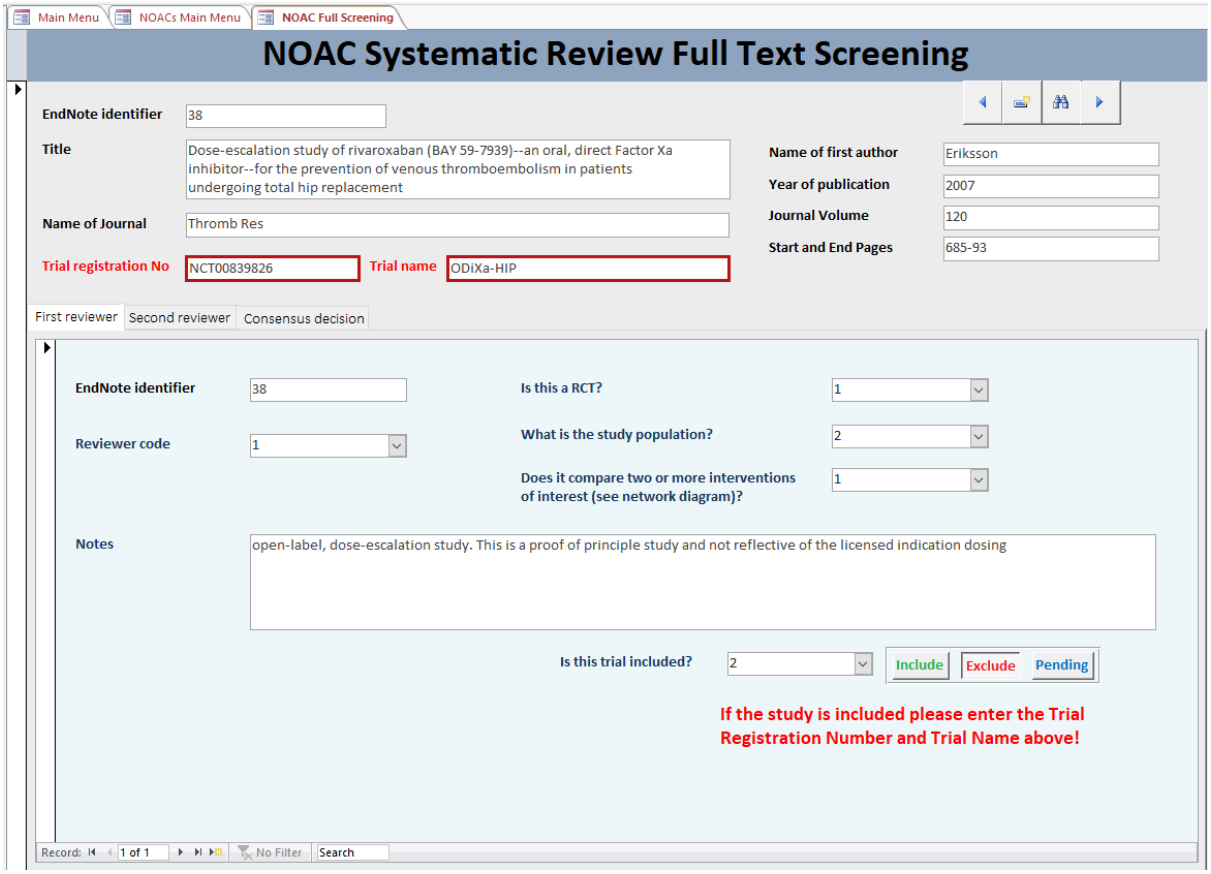


Figure 80: Screening for inclusion at full text stage

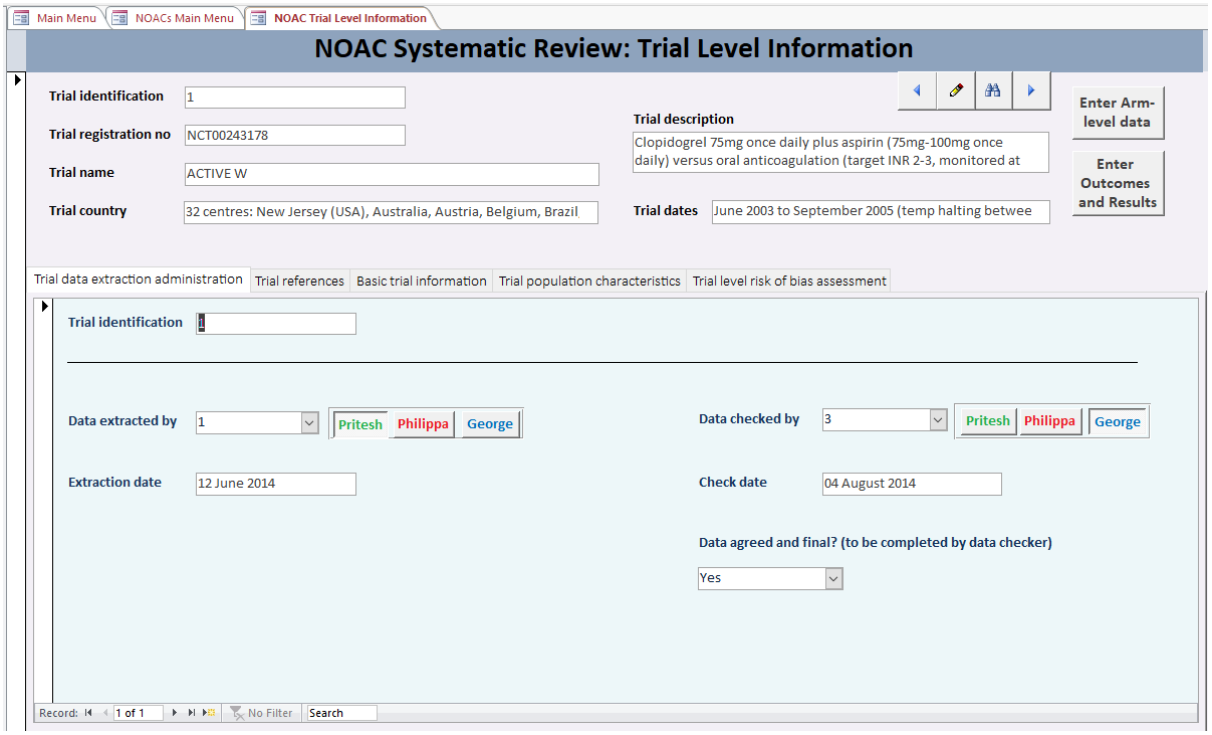


Figure 81: Trial level information (reviewer extraction build page)

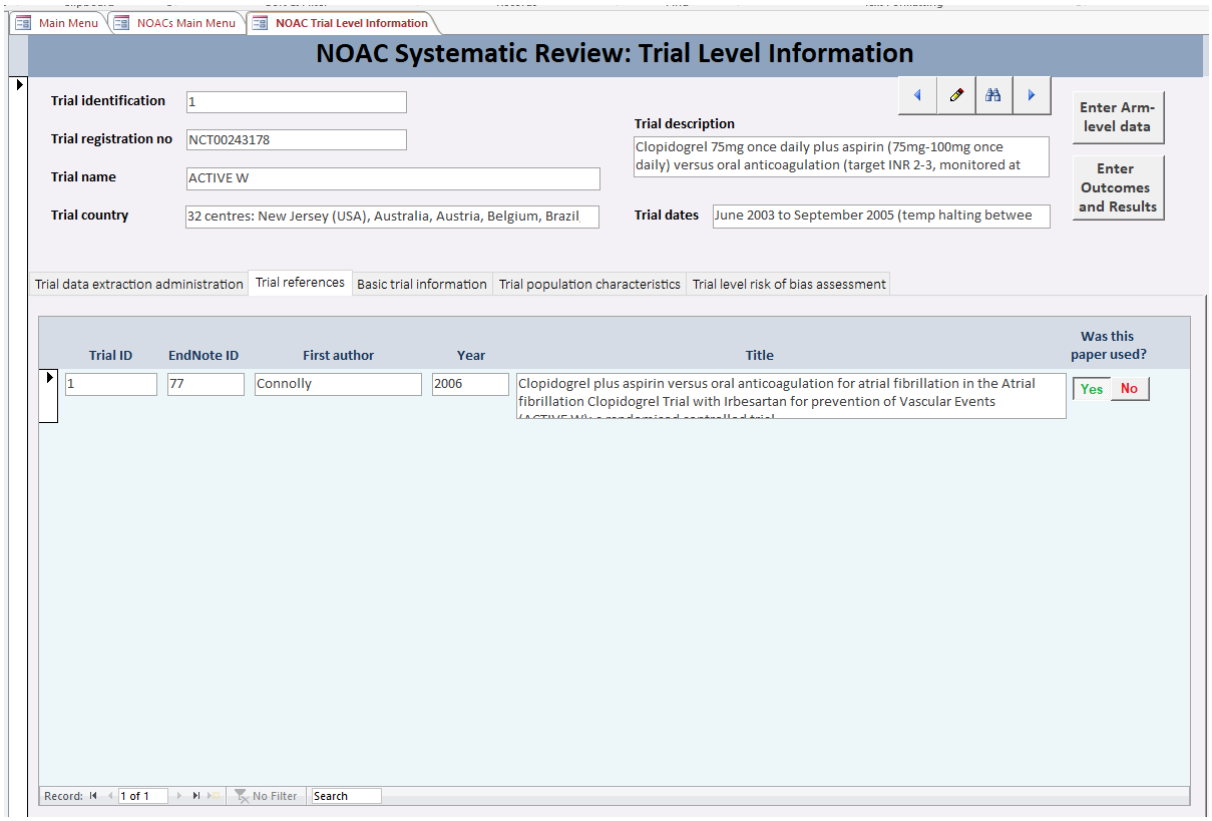


Figure 82: Trial level information (reference page)

[where more than one publication related to a single trial set of data, a TRIAL ID was created for the principal publication and all sequential ones included within this section]

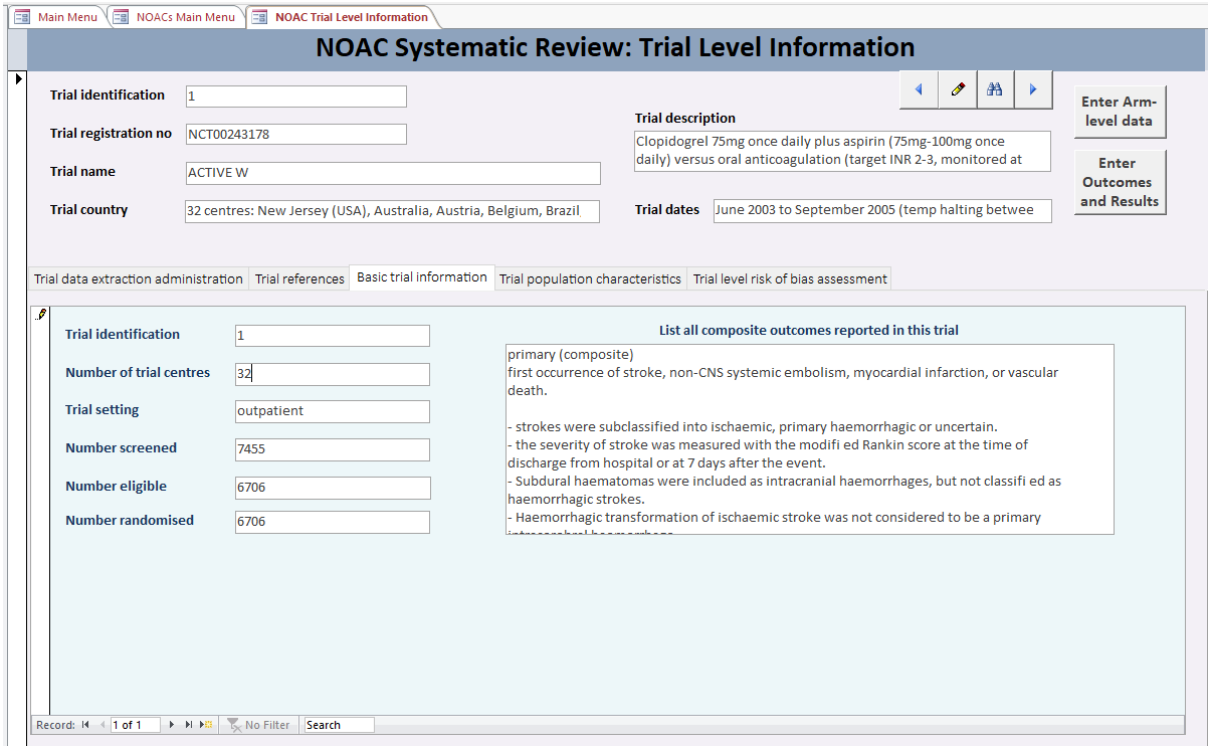


Figure 83: Trial level information (trial subject demographics and outcome information page)

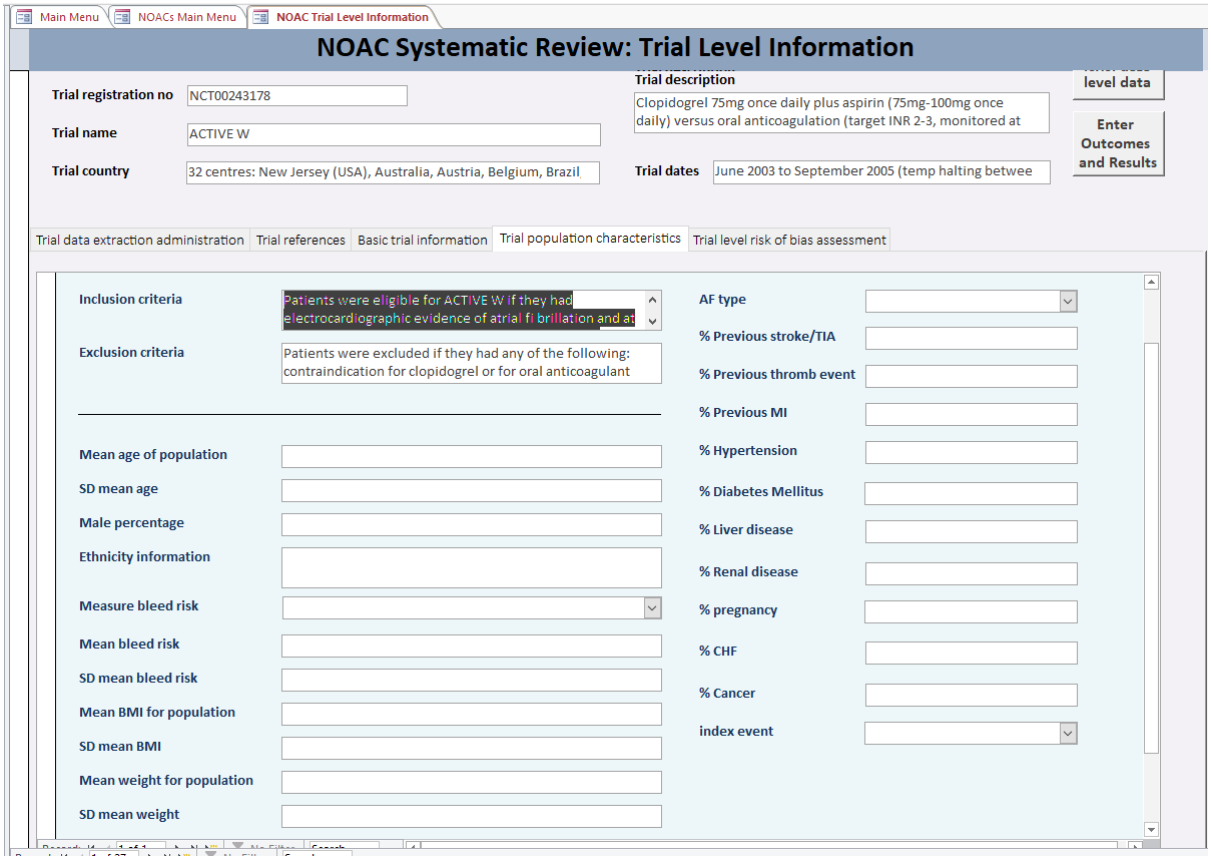


Figure 84: Trial level information (trial subject detailed characteristics page)

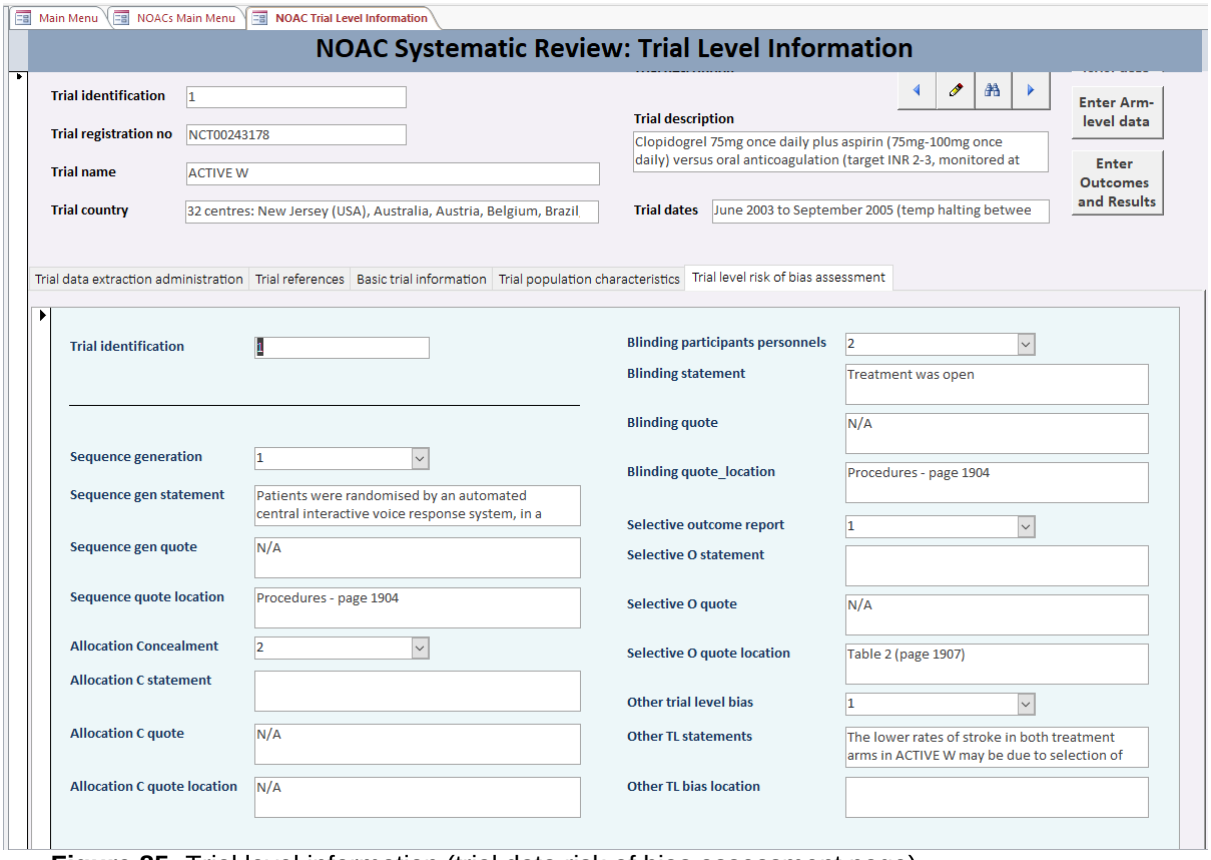


Figure 85: Trial level information (trial data risk of bias assessment page)

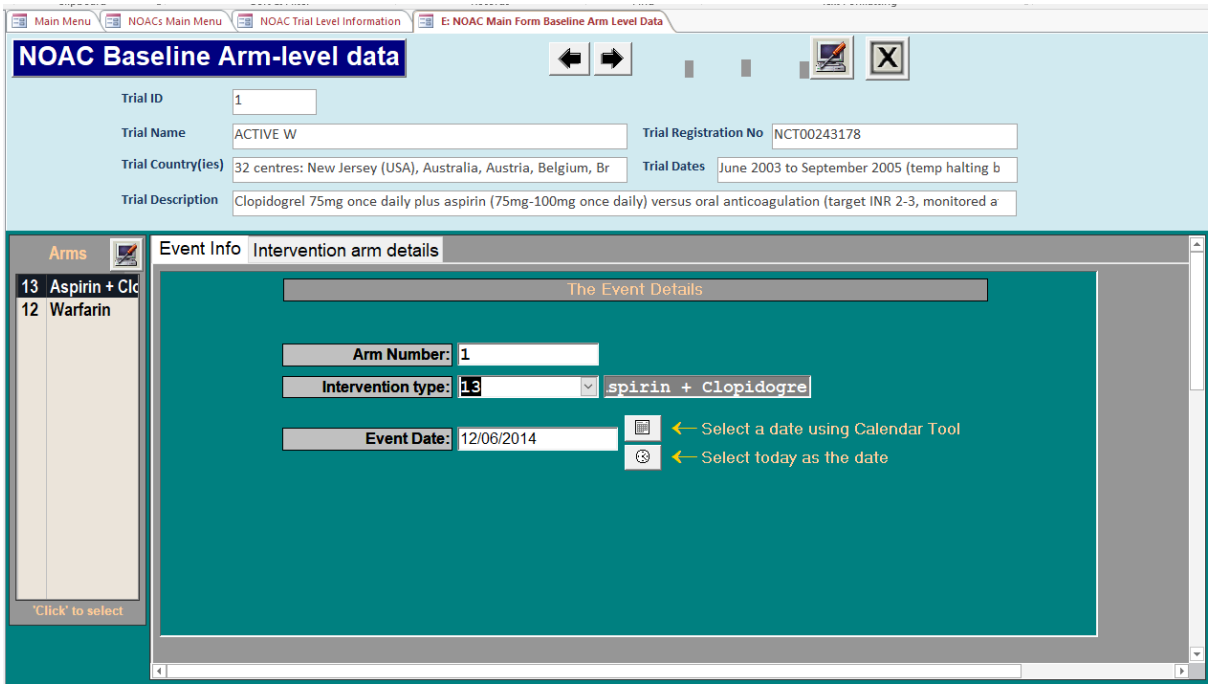


Figure 86: Trial level information (arms build page)

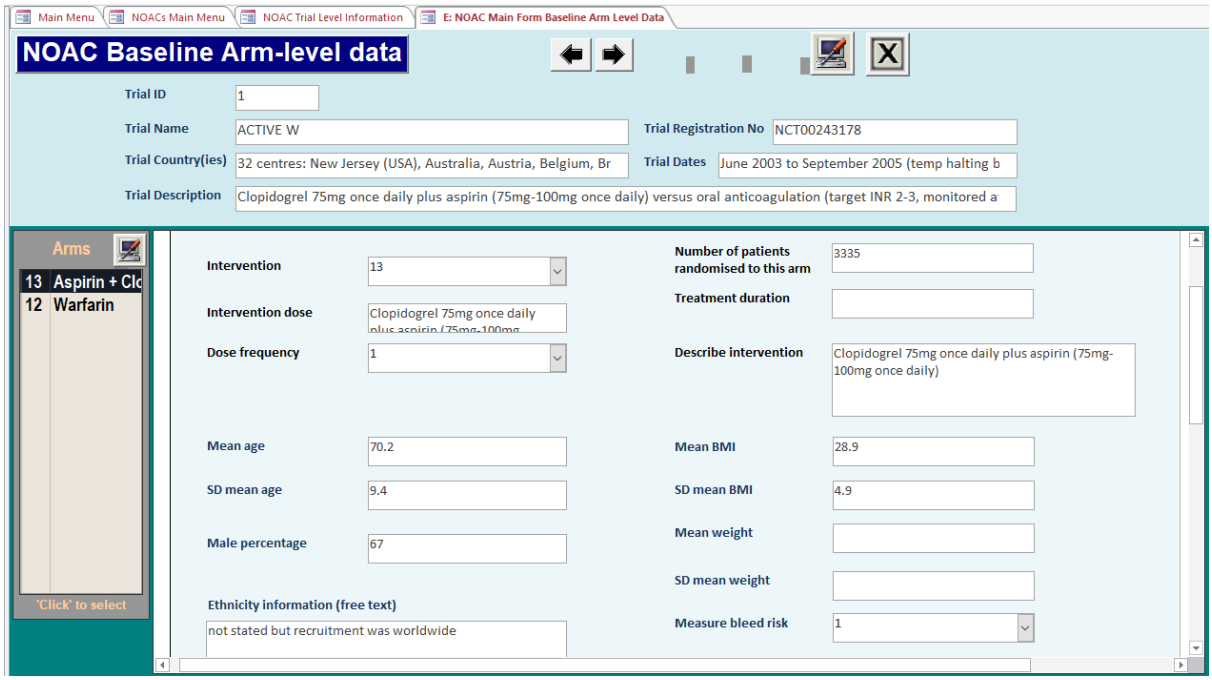


Figure 87: Trial level information (detailed arm level data page)

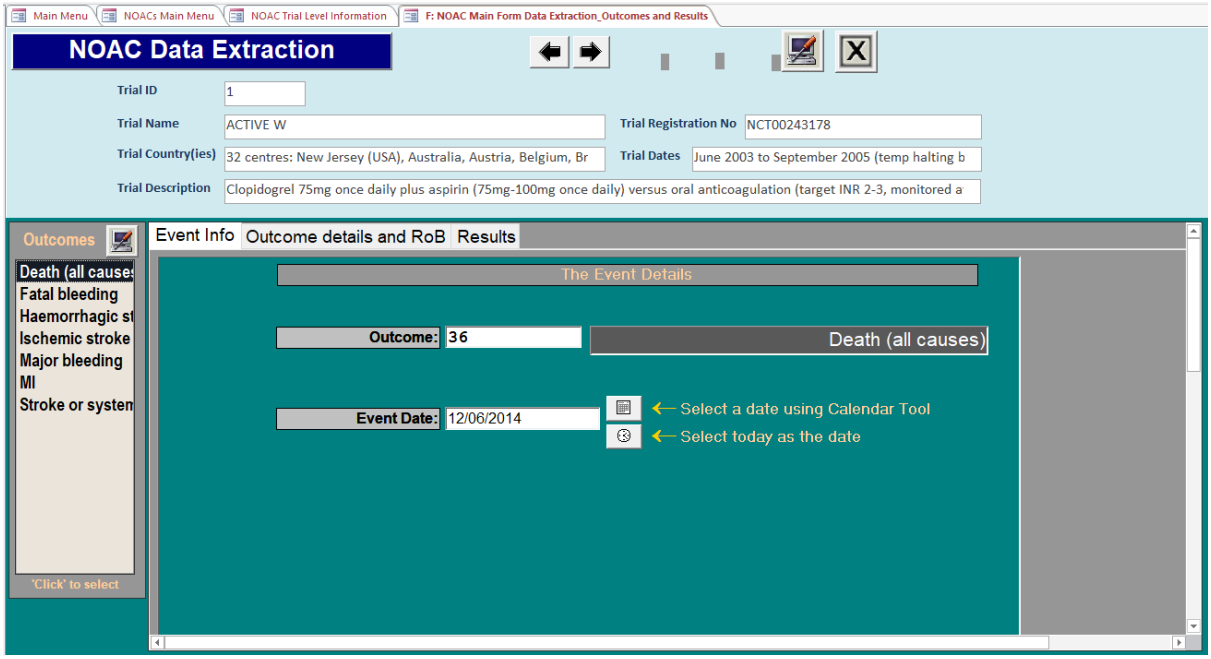


Figure 88: Trial data extraction (outcomes build page)

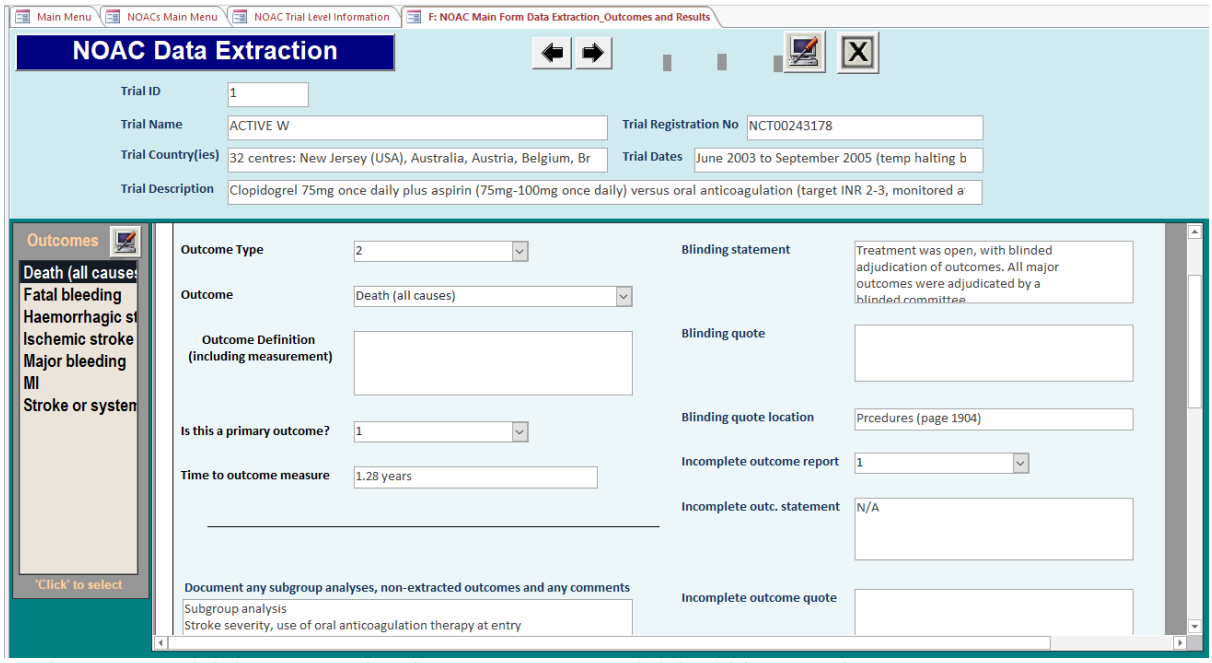


Figure 89: Trial data extraction (outcomes type and risk of bias page)

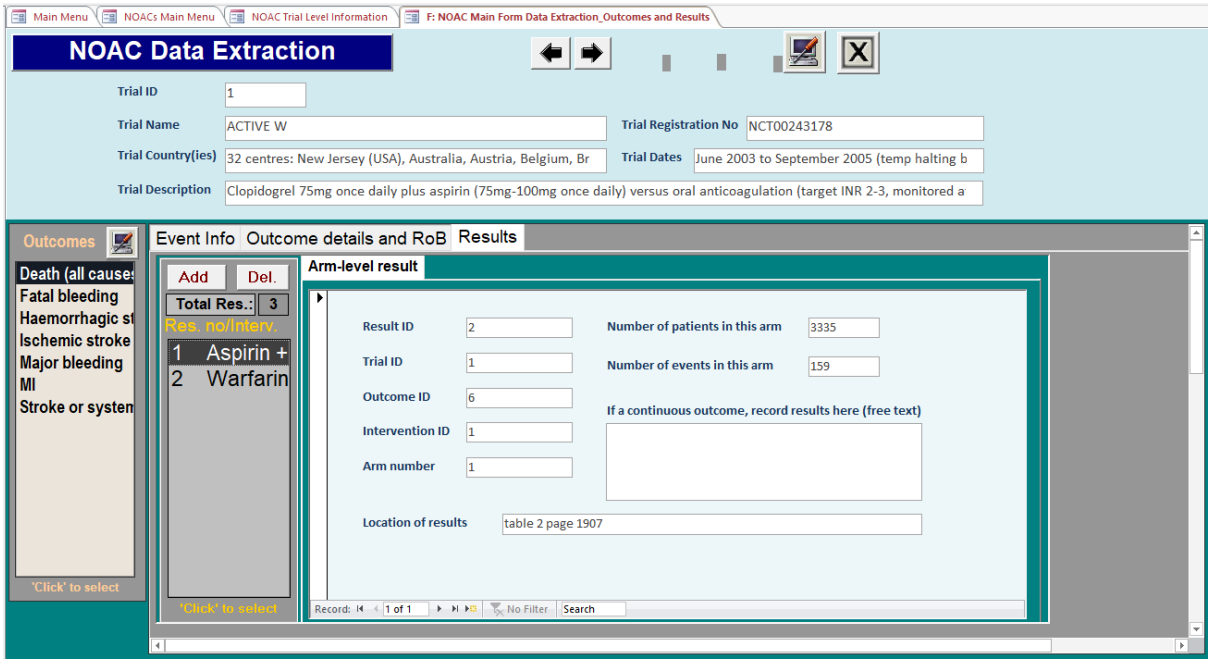


Figure 90: Trial data extraction (arm level data)

10.6 Appendix VI: Forest Plots (stroke prevention in AF)

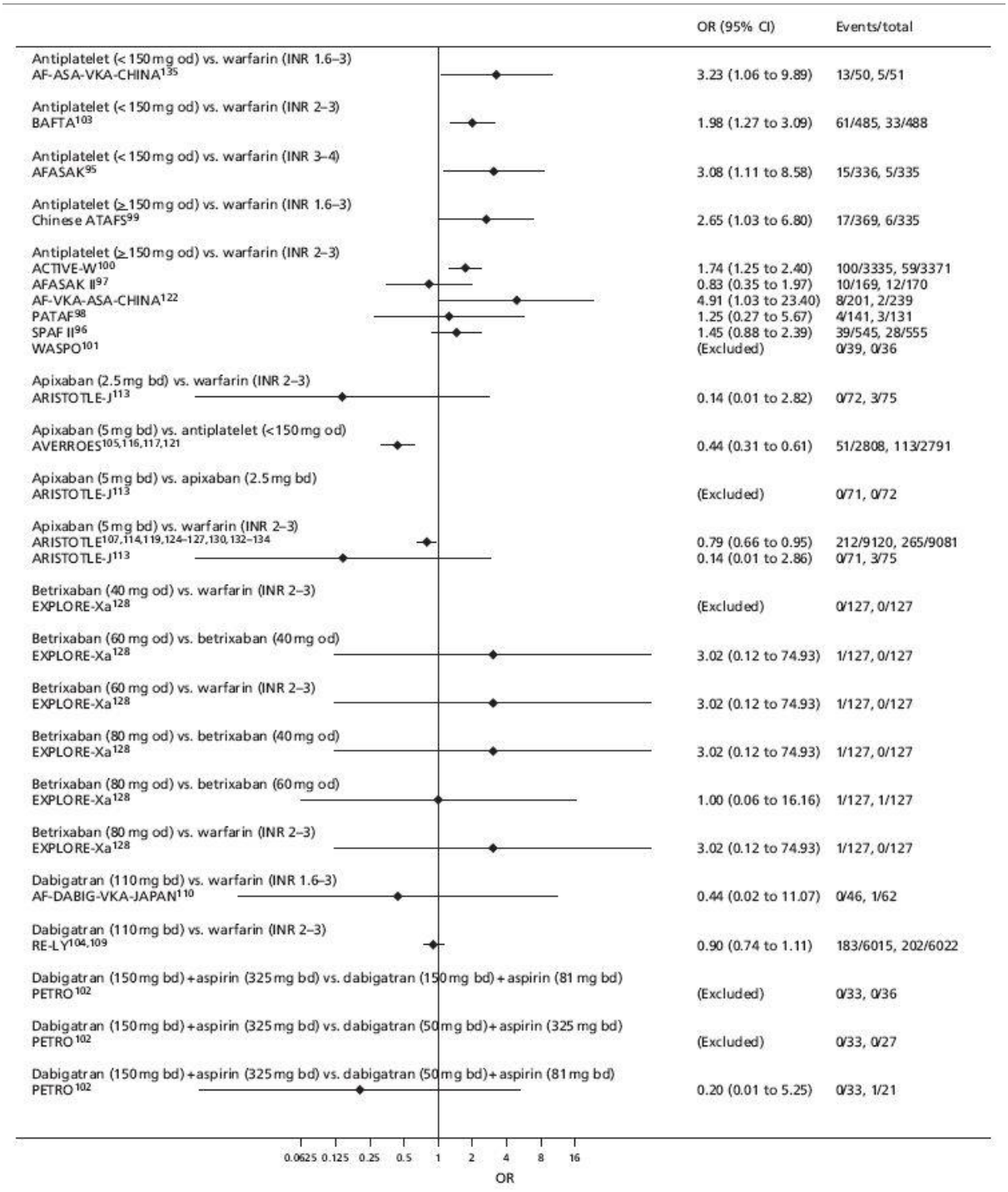


Figure 91: Forest plot of stroke or SE [1/4]

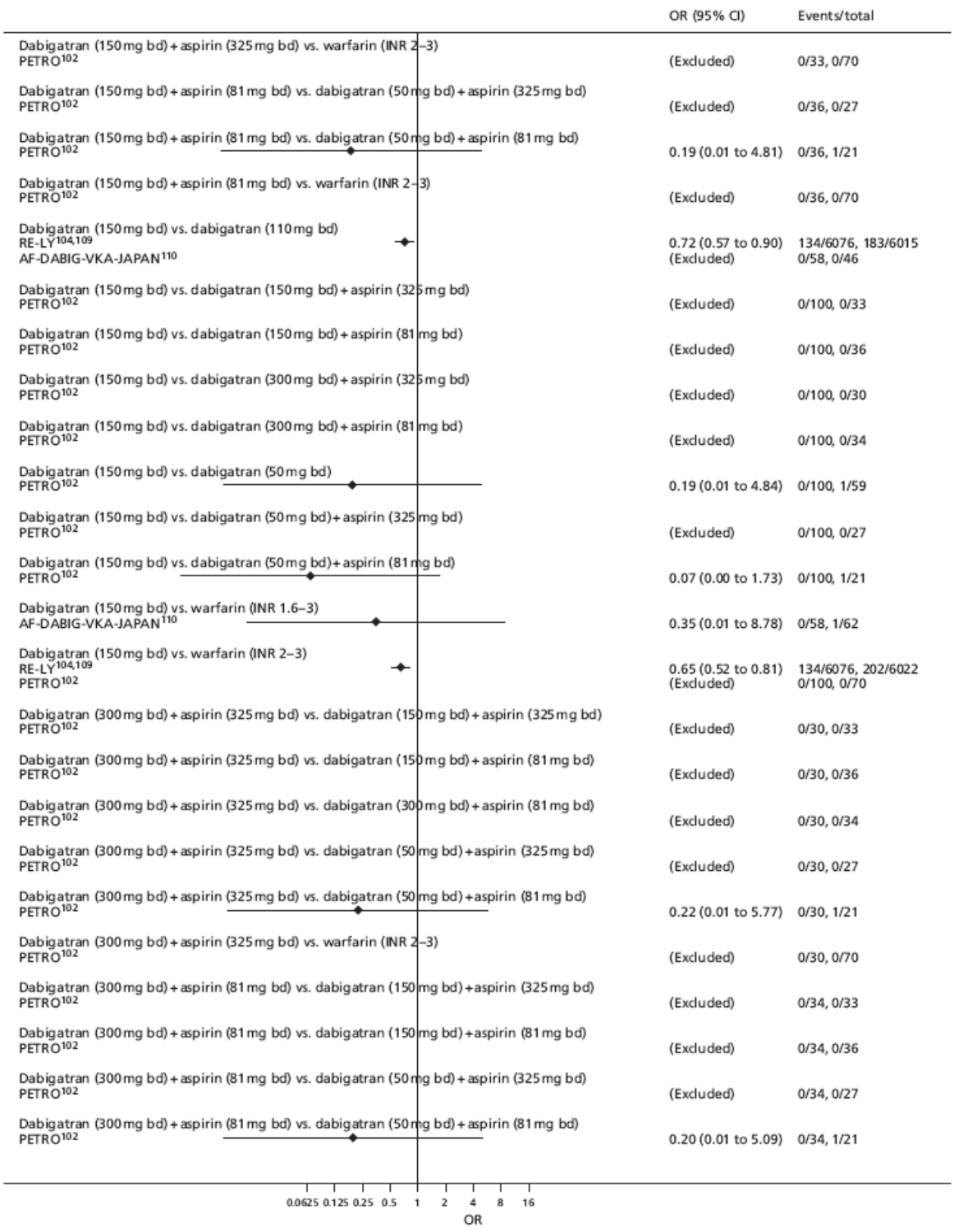


Figure 92: Forest plot of stroke or SE [2/4]

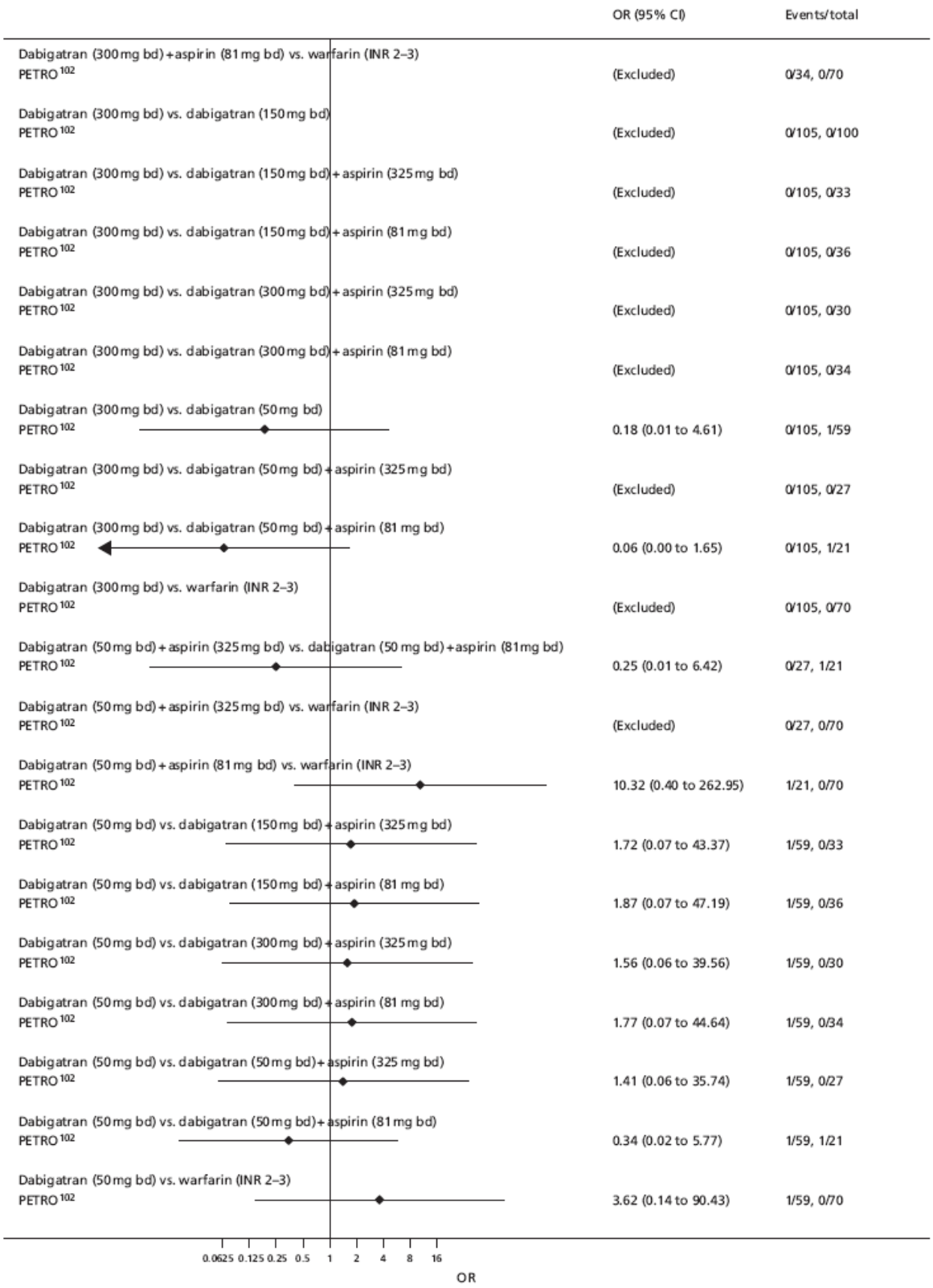


Figure 93: Forest plot of stroke or SE [3/4]

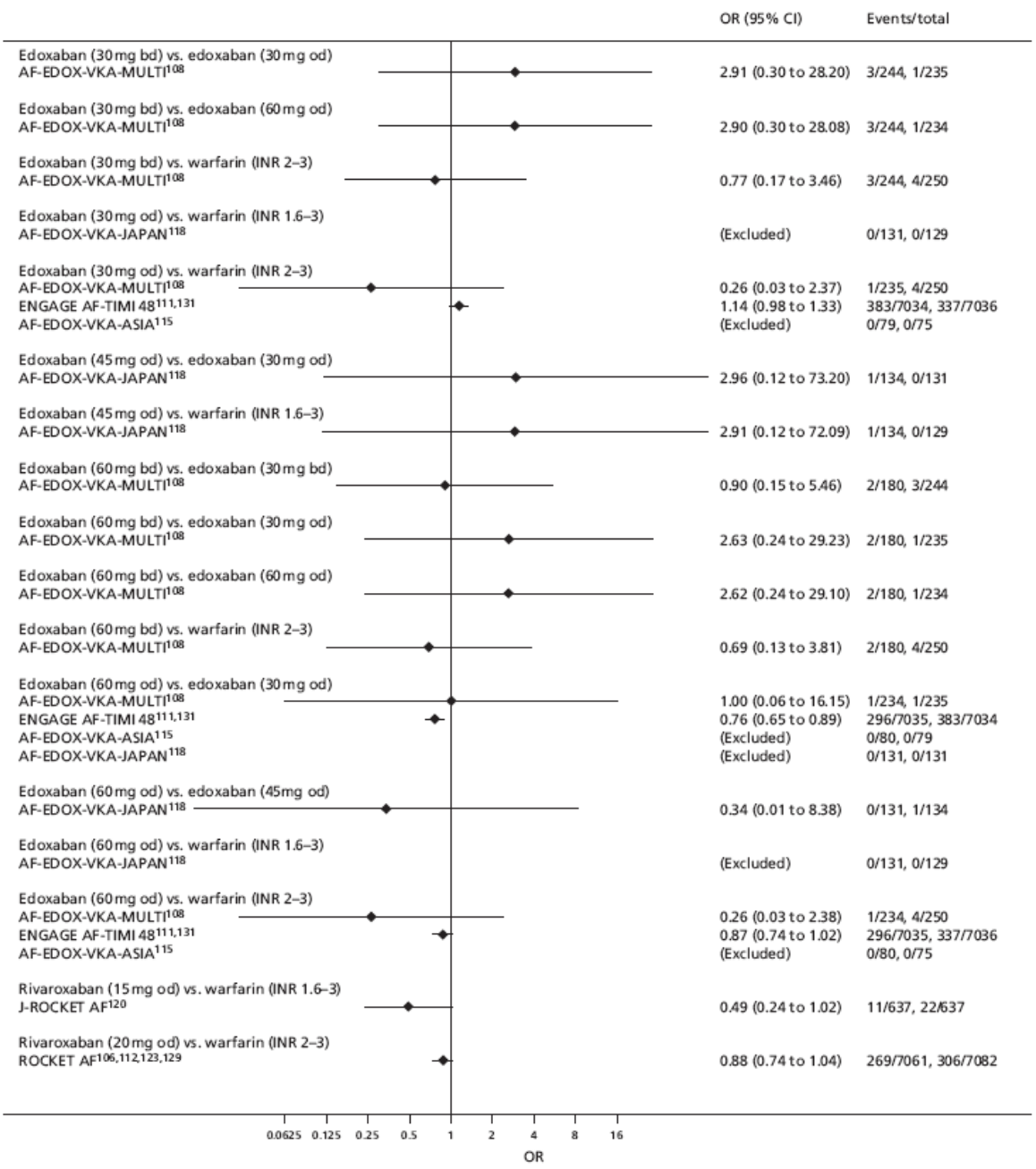


Figure 94: Forest plot of stroke or SE [4/4]

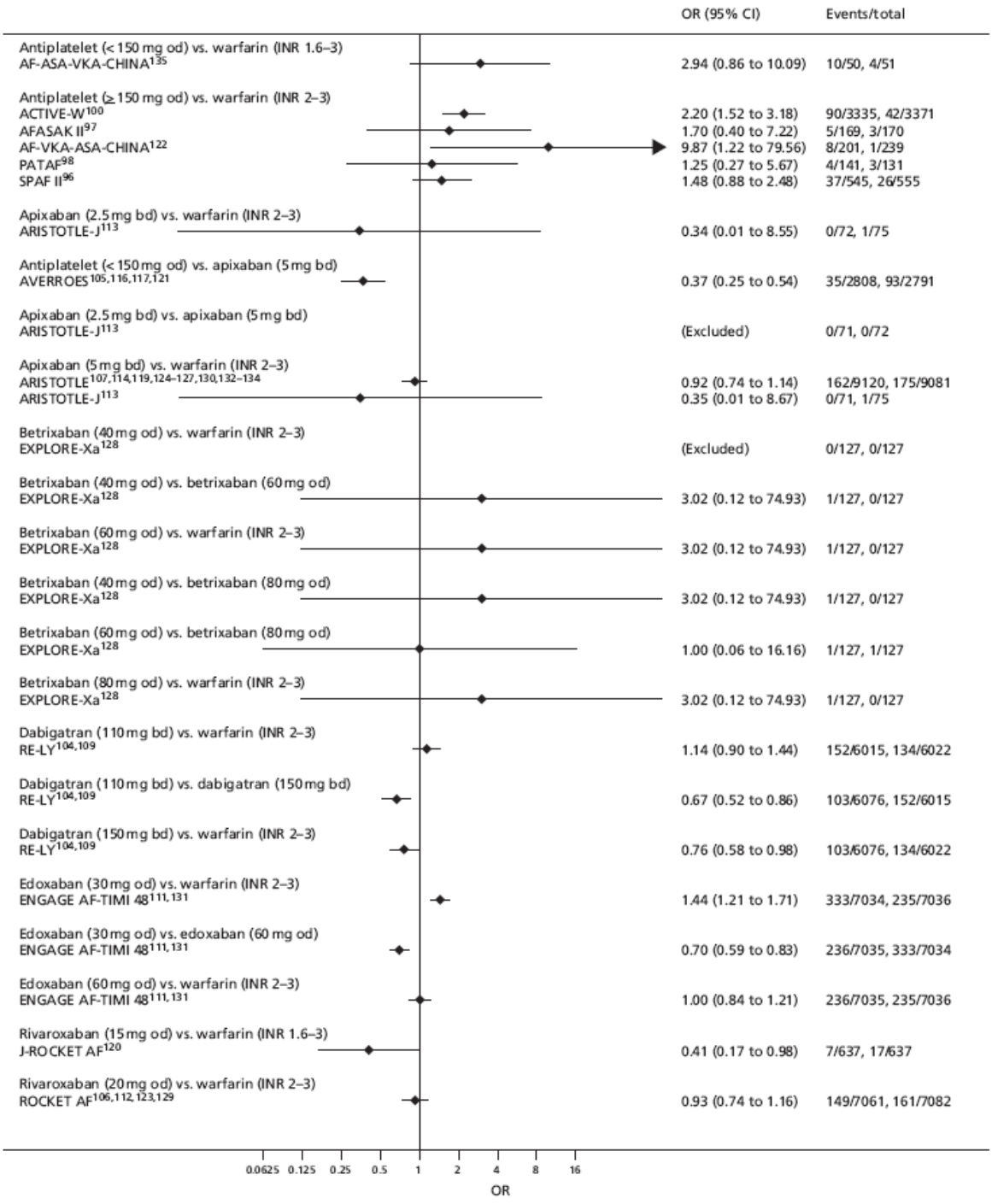


Figure 95: Forest plot of ischaemic stroke

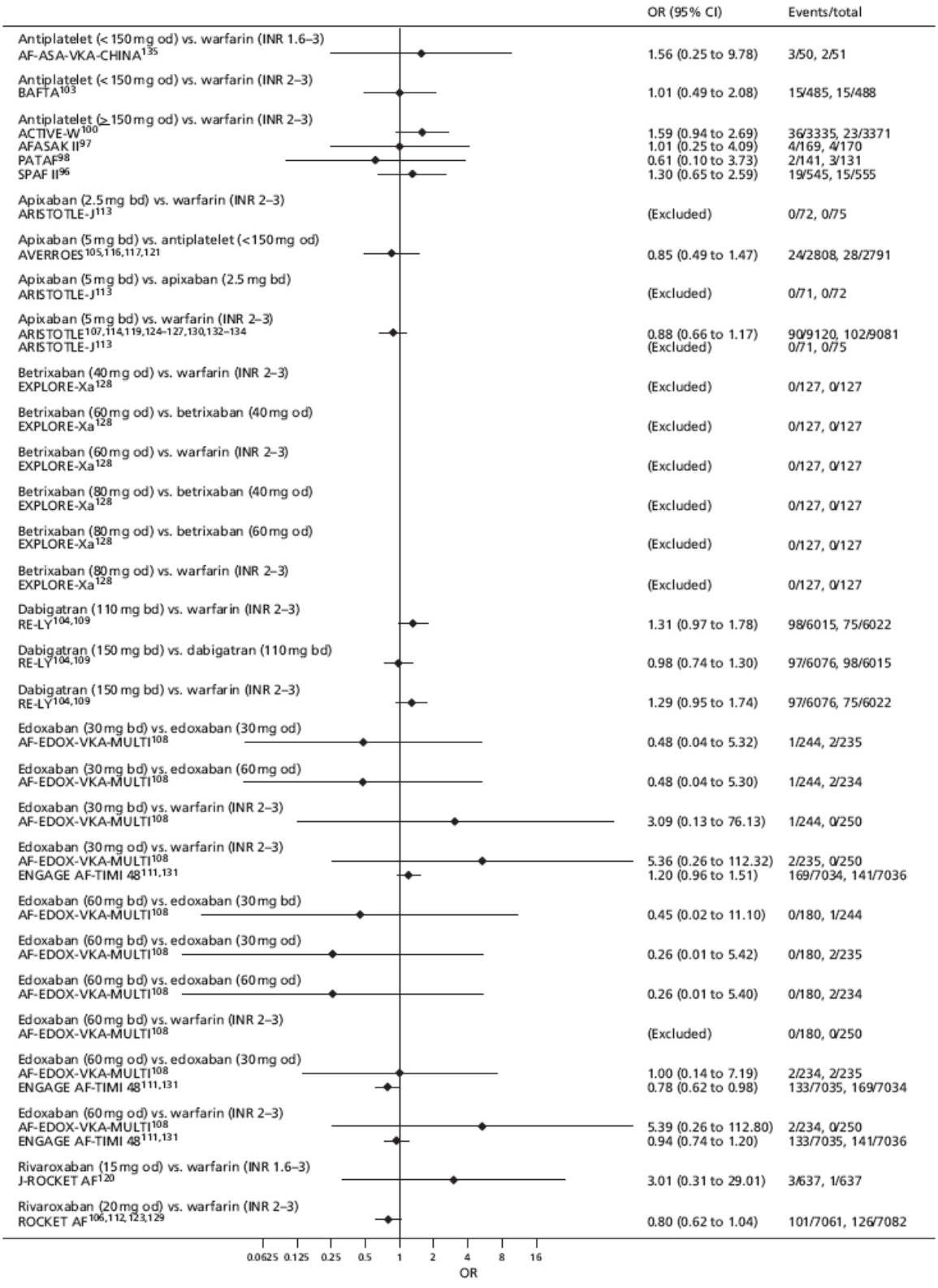


Figure 96: Forest plot of myocardial infarction

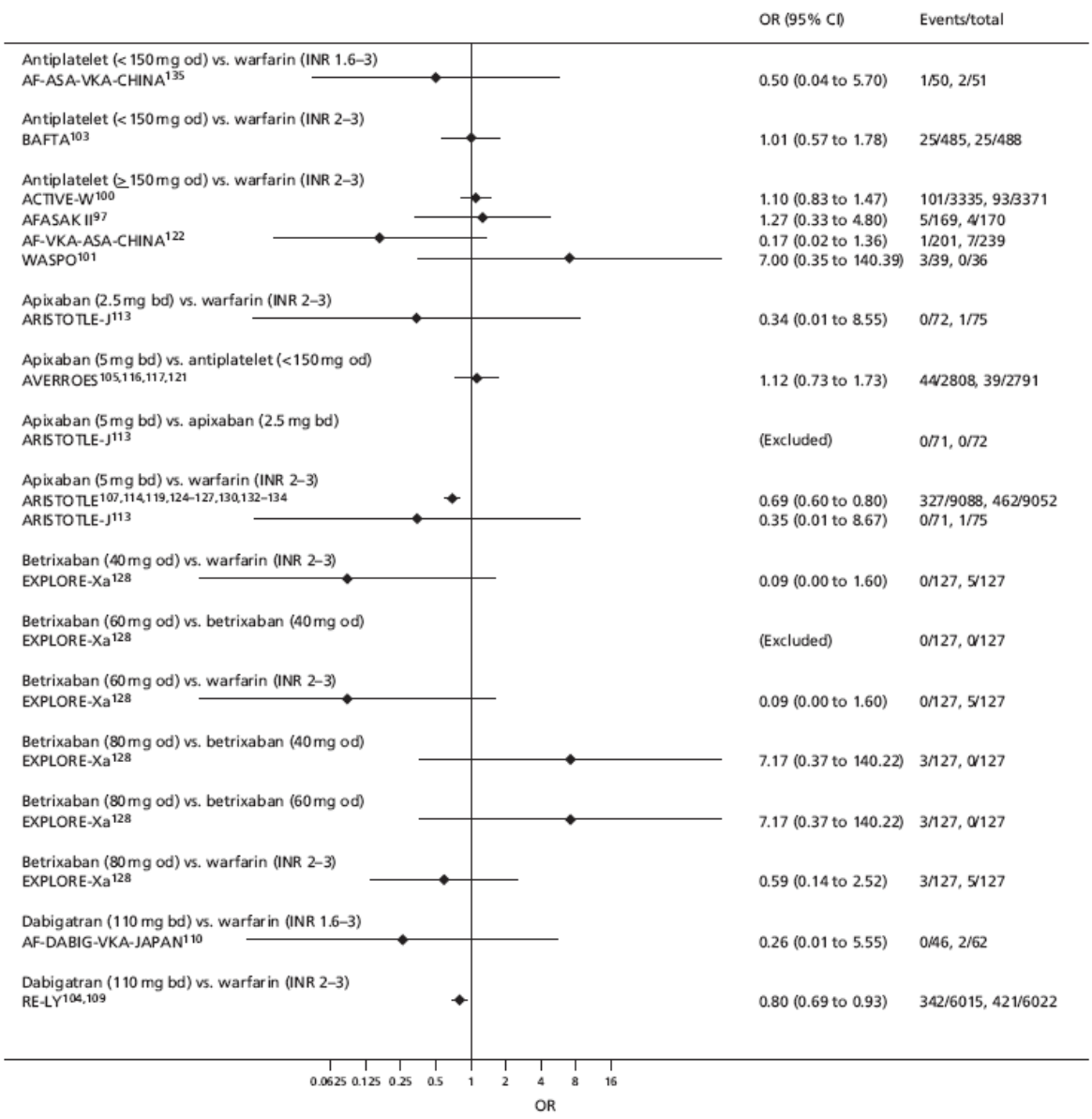


Figure 97: Forest plot of major bleeding [1/4]

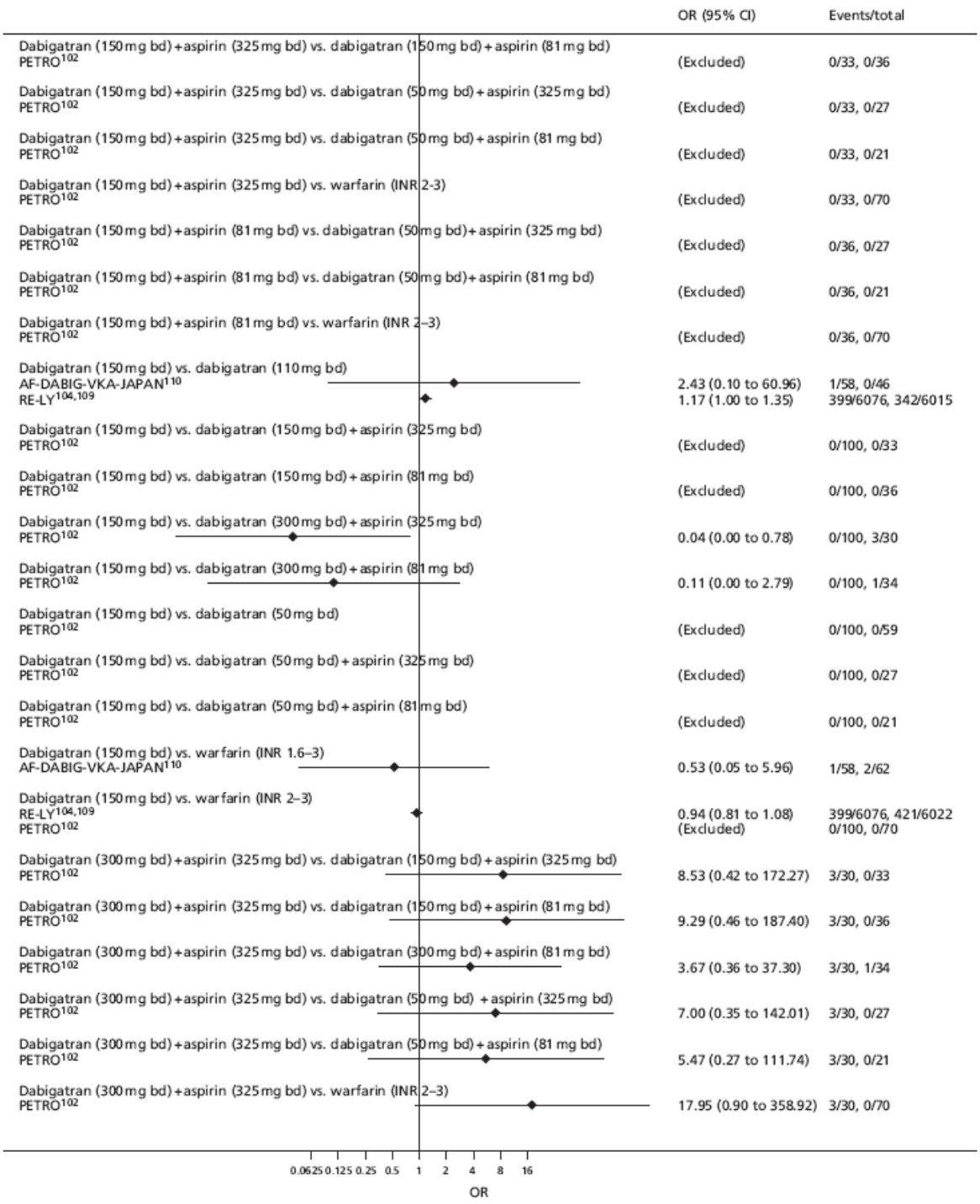


Figure 98: Forest plot of major bleeding [2/4]

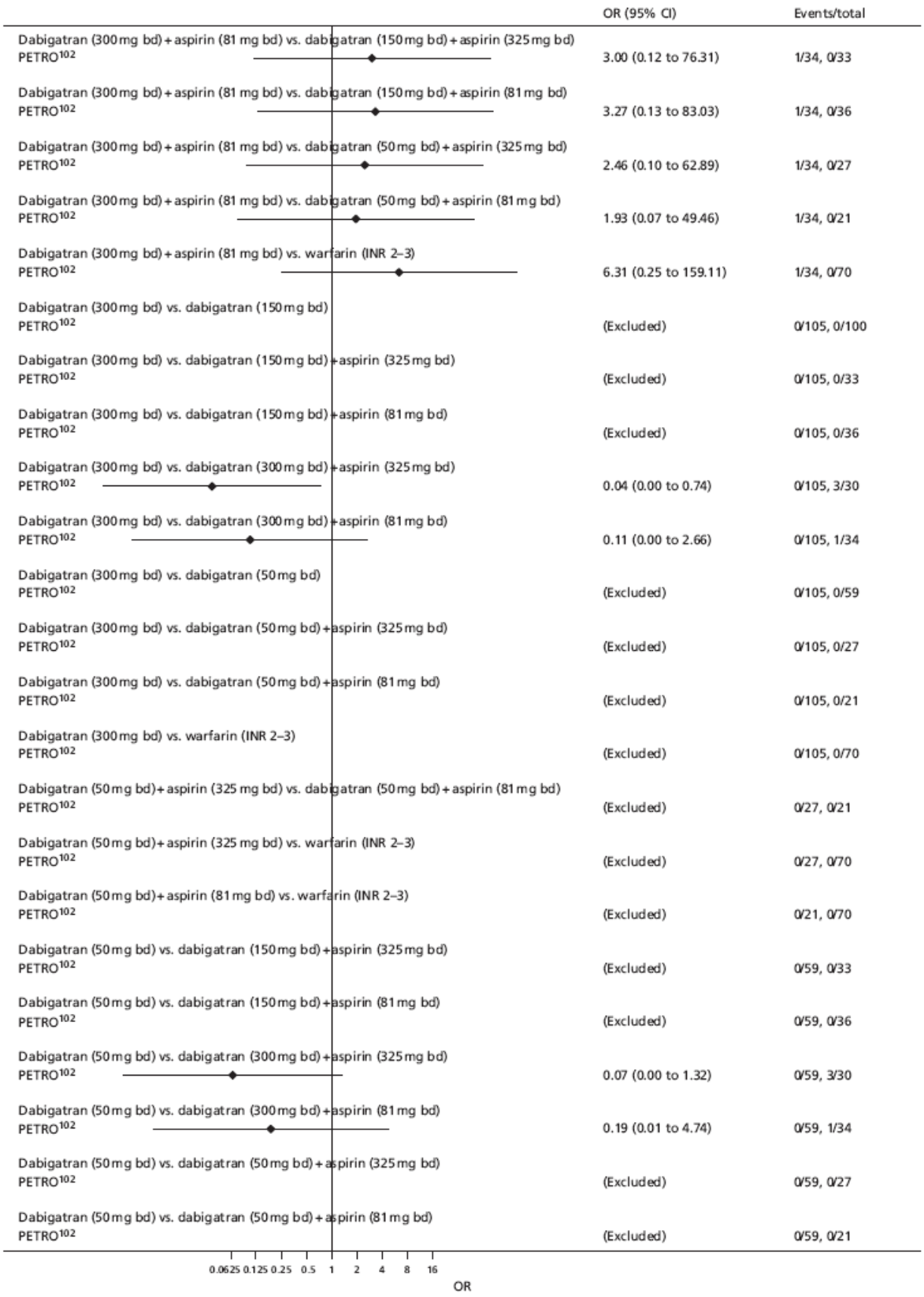


Figure 99: Forest plot of major bleeding [3/4]

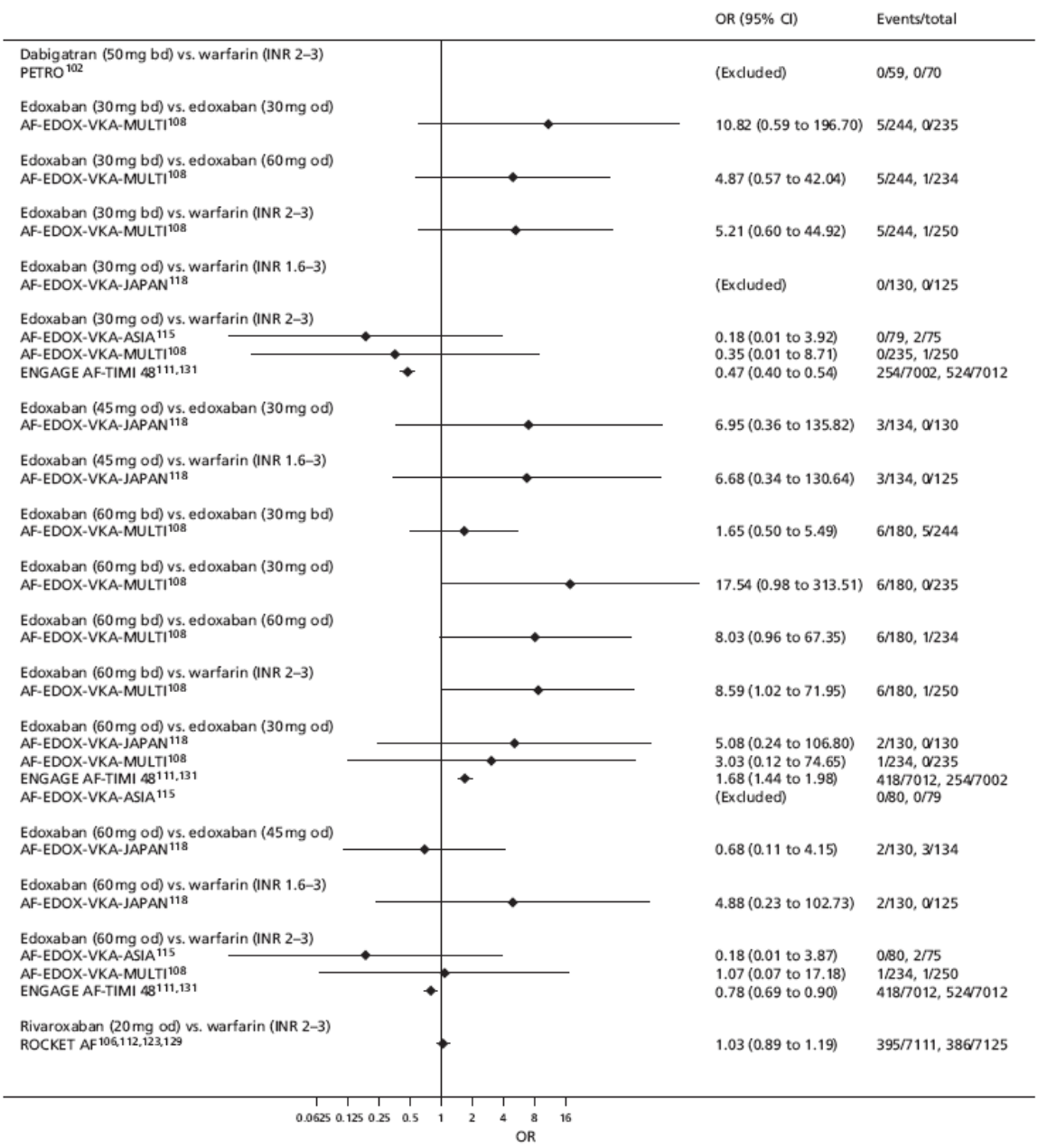


Figure 100: Forest plot of major bleeding [4/4]

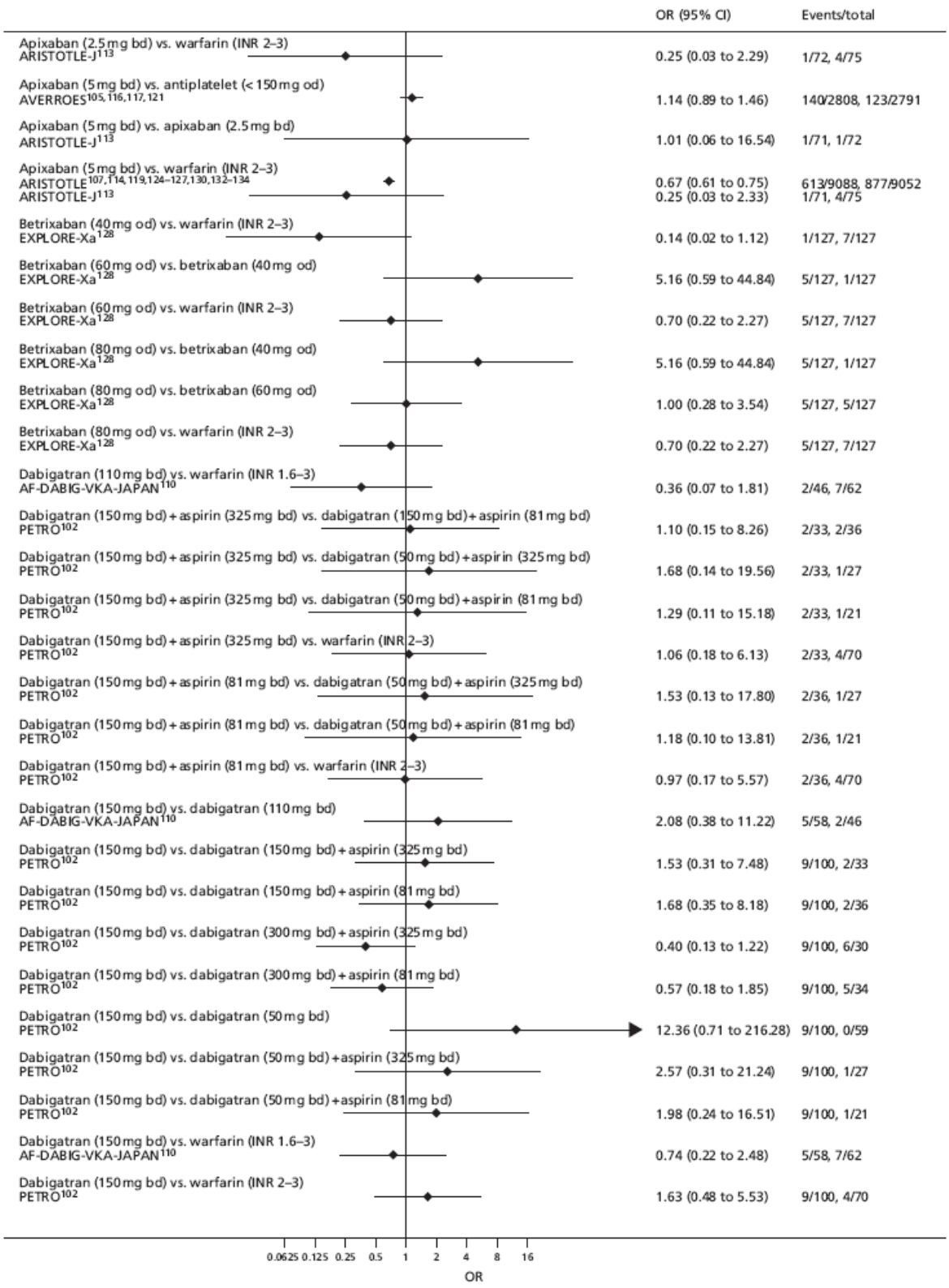


Figure 101: Forest plot of clinically relevant bleeding [1/3]

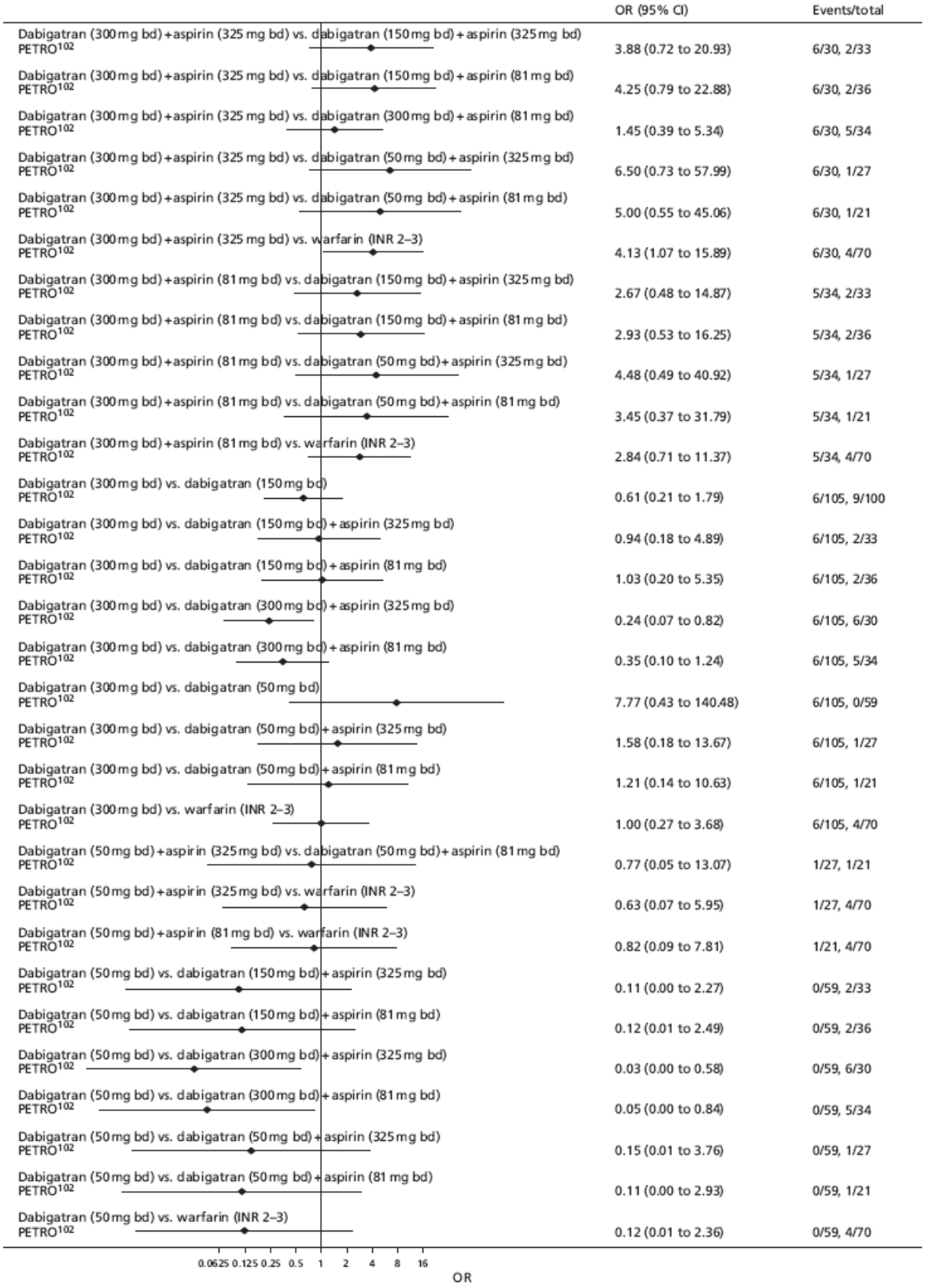


Figure 102: Forest plot of clinically relevant bleeding [2/3]

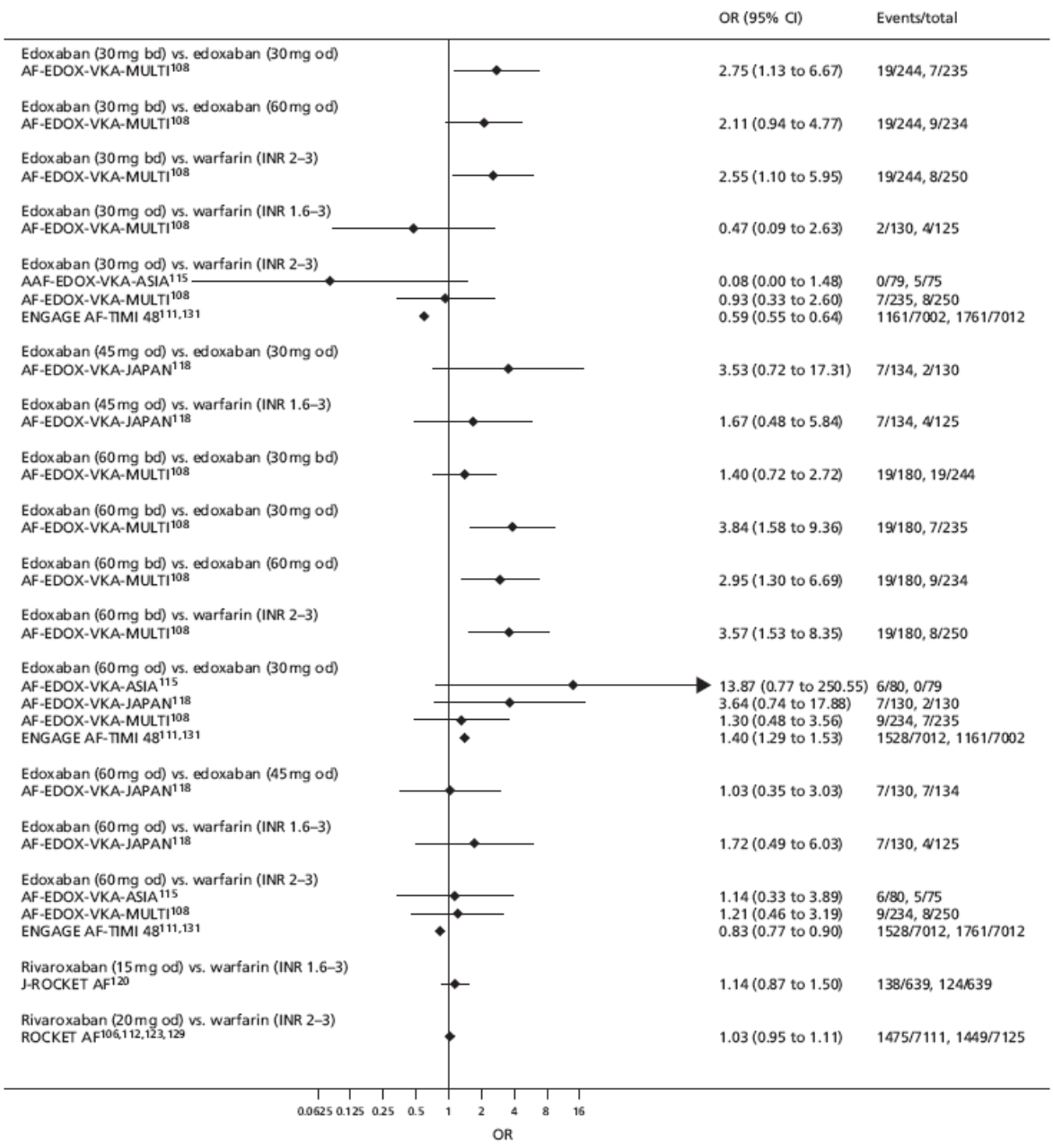


Figure 103: Forest plot of clinically relevant bleeding [3/3]

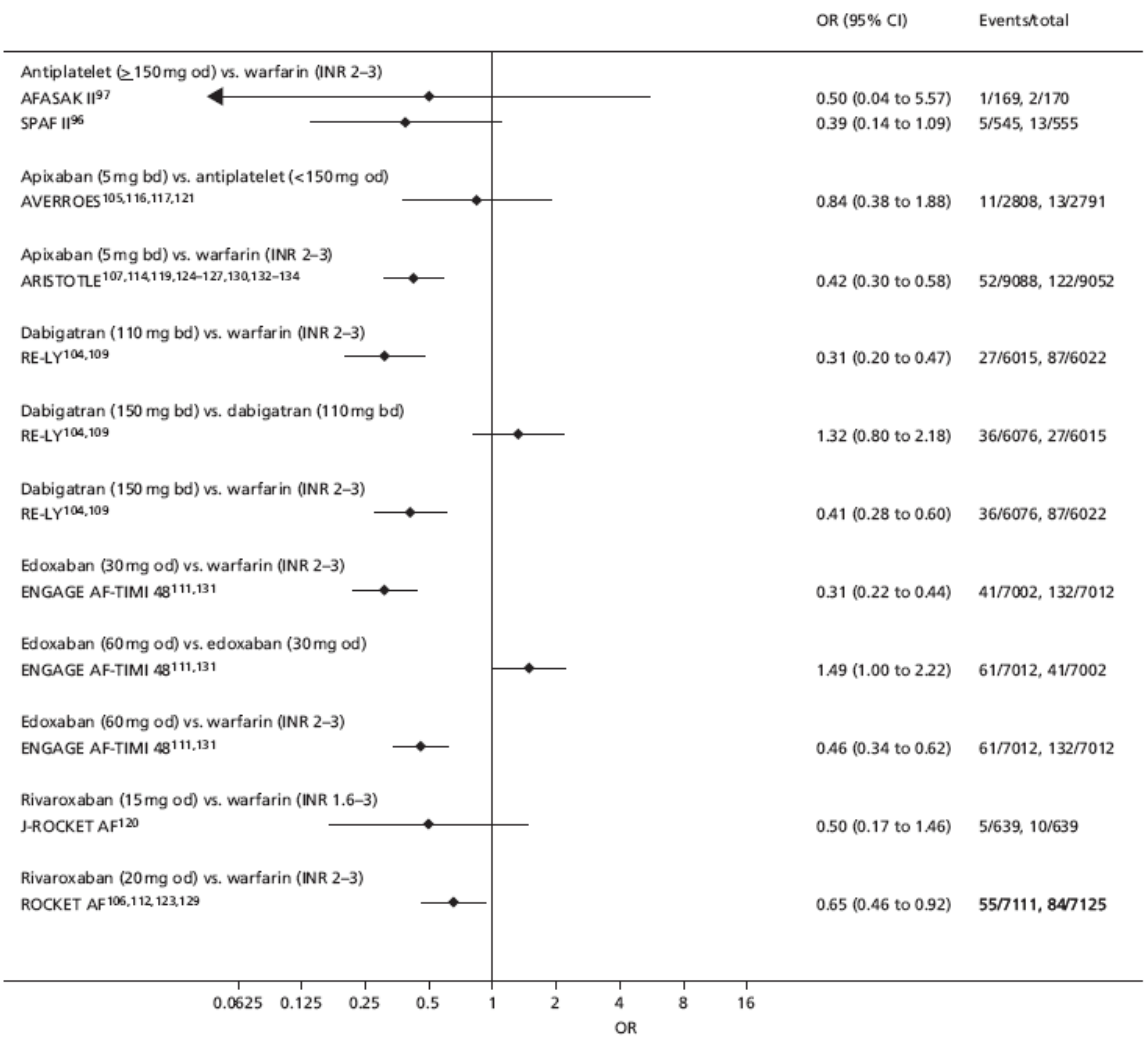


Figure 104: Forest plot of clinically intracranial bleeding

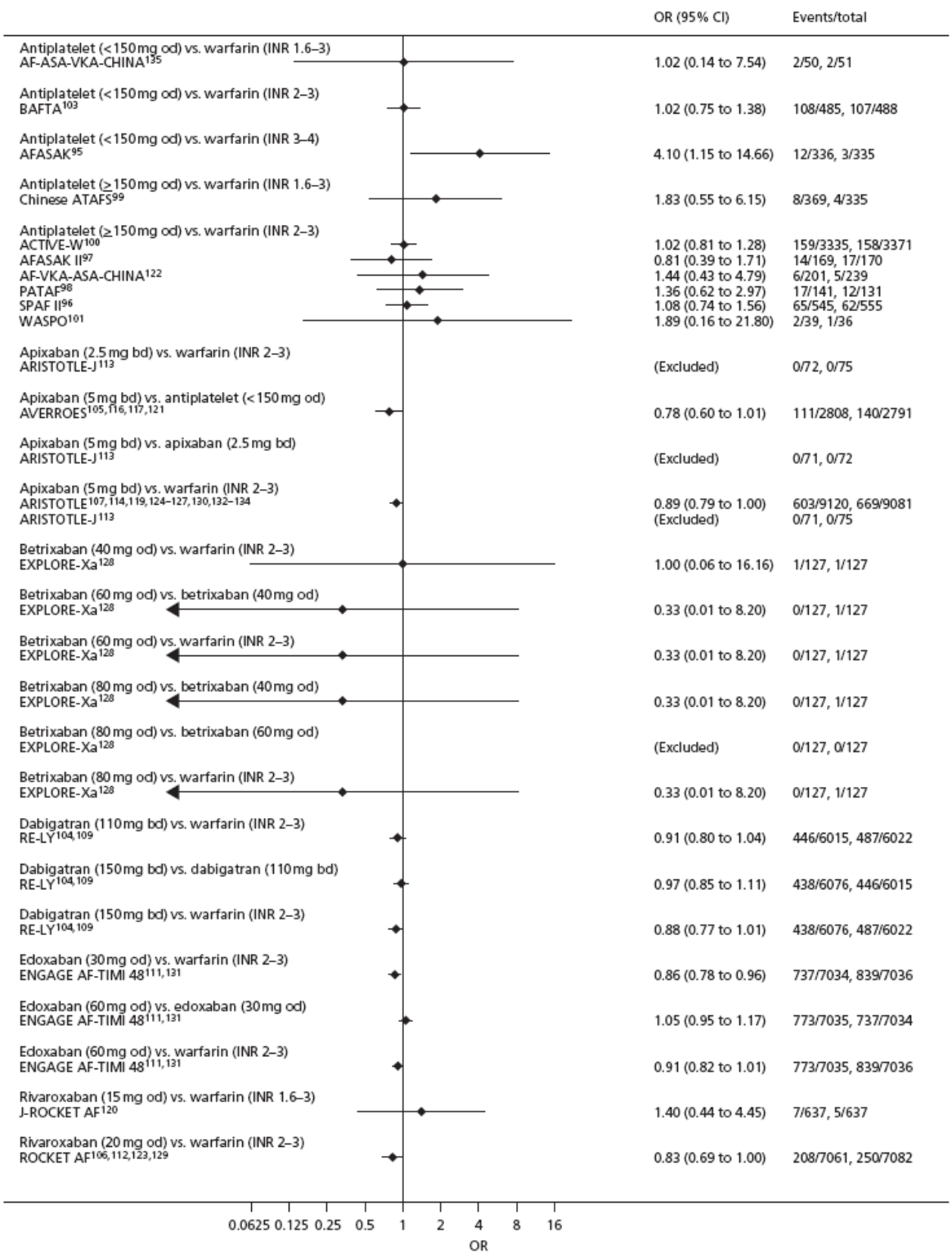


Figure 105: Forest plot of all cause mortality

10.7 Appendix VII: Antiepileptic drugs for refractory epilepsy (publication abstract)

Br J Clin Pharmacol. 2013 Nov; 76[5]: 649-667.

Bodalia PN, Grosso AM, Sofat R, MacAllister RJ, Smeeth L, Dhillon S, Casas JP, Wonderling D, Hingorani AD.

Department of Pharmacy, University College London Hospitals NHS Foundation Trust, London, UK. pritesh.bodalia@uclh.nhs.uk

AIMS: To evaluate the comparative efficacy (50% reduction in seizure frequency) and tolerability (premature withdrawal due to adverse events) of anti-epileptic drugs (AEDs) for refractory epilepsy.

METHODS: We searched Cochrane Central Register of Controlled Trials (Cochrane Library 2009, issue 2) including Epilepsy Group's specialized register, MEDLINE (1950 to March 2009), EMBASE (1980 to March 2009), and Current Contents Connect (1998 to March 2009) to conduct a systematic review of published studies, developed a treatment network and undertook a network meta-analysis.

RESULTS: Forty-three eligible trials with 6346 patients and 12 interventions, including placebo, contributed to the analysis. Only three direct drug comparator trials were identified, the remaining 40 trials being placebo-controlled. Conventional random-effects meta-analysis indicated all drugs were superior in efficacy to placebo (overall odds ratio (OR] 3.78, 95% CI 3.14, 4.55) but did not permit firm distinction between drugs on the basis of the efficacy or tolerability. A Bayesian network meta-analysis prioritized oxcarbazepine, topiramate and pregabalin on the basis of short term efficacy. However, sodium valproate, levetiracetam, gabapentin and vigabatrin were prioritized on the basis of short-term efficacy and tolerability, with the caveat that vigabatrin is recognized as being associated with serious visual disturbance with chronic use.

CONCLUSION: Of the wide range of AEDs licensed for the treatment of refractory epilepsy, sodium valproate, levetiracetam and gabapentin demonstrated the best balance of efficacy and tolerability. Until regulators mandate greater use of active comparator trials with longer term follow-up, network meta-analysis provides the only available means to quantify these clinically important parameters.

10.8 Appendix VIII: Angiotensin Receptor Blockers for the management of hypertension and heart failure (publication abstract)

Int J Clin Practice 2011 March; 65[3]: 253-263.

Grosso AM, Bodalia PN, MacAllister RJ, Hingorani AD, Moon JC, Scott MA.

Department of Pharmacy, University College London Hospitals NHS Foundation Trust, London, UK. pritesh.bodalia@uclh.nhs.uk

BACKGROUND: The UK National Health Service (NHS) currently spends in excess of £250 million per annum on angiotensin II receptor blockers (ARBs) for the treatment of hypertension and heart failure; with candesartan currently dominating the market. With the recent introduction of generic losartan, we set out to directly compare the branded market leader to its now cheaper alternative.

AIMS: The primary objectives were to compare the blood pressure (BP) lowering efficacy and cardiovascular outcomes of candesartan and losartan in the treatment of essential hypertension and chronic heart failure, respectively. The secondary objective was to model their comparative incremental cost-effectiveness in a UK NHS setting.

METHODS: The Cochrane Central Register of Controlled Trials (*Cochrane Library* 2009, issue 2), which contains the Hypertension and Heart Group's specialist register, Medline (1950–February 2010), and Embase (1980–February 2010) were included in the search strategy. Selection criteria were randomised studies of candesartan versus losartan in adults (> 18 years). The main outcome measures were as follows: Hypertension: mean change from baseline in trough (24 h postdose) systolic and diastolic BP. Heart failure: composite of cardiovascular death and hospital admission for management of heart failure.

RESULTS: Two reviewers applied inclusion criteria, assessed trial quality, and extracted data. Eight (three of which met inclusion criteria) and zero trials compared candesartan directly with losartan in the treatment of hypertension and heart failure, respectively. A between-treatment difference of −1.96 mmHg [95% confidence interval (CI) −2.40 to −1.51] for trough diastolic BP and −3.00 mmHg (95% CI −3.79 to −2.22) for trough systolic BP in favour of candesartan was observed. Based on this differential, a 10-year Markov model estimates the cost per quality-adjusted life-year gained to exceed £40,000 for using candesartan in place of generic losartan.

CONCLUSION: Candesartan reduces BP to a slightly greater extent when compared with losartan; however, such difference is unlikely to be cost-effective based on current acquisition costs, perceived NHS affordability thresholds and use of combination regimens. We could find no robust evidence supporting the superiority of candesartan over losartan in the treatment of heart failure. We therefore recommend using generic losartan as the ARB of choice which could save the UK NHS approximately £200 million per annum in drug costs.

10.9 Appendix IX: Antimuscarinics for the management of overactive bladder syndrome (abstract)

DRAFT abstract [unpublished]

Bodalia PN, Wood DN, Welton NJ.

Department of Pharmacy, University College London Hospitals NHS Foundation Trust, London, UK. pritesh.bodalia@uclh.nhs.uk

BACKGROUND: Guidelines for the treatment of Overactive Bladder (OAB) currently recommend antimuscarinics as first-line pharmacological therapy. Mirabegron is then considered for subjects in whom an antimuscarinic is contraindicated, clinically ineffective or intolerable. However, the optimal choice for frontline antimuscarinic therapy remains unclear with a variety of licensed agents available.

AIM: The aim of this study was to perform a mixed-treatment comparison (MTC) to define a prescribing hierarchy for antimuscarinics in OAB.

METHODS: We searched Cochrane Central Register of Controlled Trials (Cochrane Library 2009, issue 2), MEDLINE (1950 to June 2014) and EMBASE (1980 to June 2014) to conduct a systematic review of published studies, developed a treatment network and undertake a network meta-analysis. Selection criteria were randomised studies of pharmacological treatments in the management of OAB in adults (> 18 years) that were published after year 2000 to supplement new studies following a previous systematic review. The main outcome measures were as follows: micturition frequency per 24 hours; urgency urinary incontinence (UUI) episodes per 24 hours; incidence of dry mouth; and incidence of constipation.

RESULTS: The MTC analysis on micturition frequency was based on 21 studies involving 18,863 patients. The median placebo-corrected difference from baseline to week 12 was statistically significantly different from placebo for all treatments except darifenacin. Solifenacin 10mg was deemed the most effective agent for the ranking analysis. The MTC analysis on urge urinary incontinence (UUI) episodes was based on 17 studies involving 15,502 patients). The median placebo-corrected difference from baseline to week 12 was statistically significantly different for all treatments except immediate-release 4mg tolterodine, 4mg fesoterodine, 40mg immediate-release trospium and darifenacin. 10mg Solifenacin demonstrated the greatest reduction in number of UUI episodes with 5mg solifenacin and 10mg modified-release oxybutynin demonstrating joint second greatest reduction.

CONCLUSION: All antimuscarinics were associated with recognised anticholinergic adverse events such as dry mouth and constipation. In general, treatment efficacy appears to be inversely proportional to incidence of adverse events with the possible exception of darifenacin, which appeared to perform sub-optimally in terms of both efficacy and safety.

10.10 Appendix X: Novel Oral Anticoagulants for the prevention of stroke in AF (abstract)

DRAFT submission to BMJ

Sterne JAC, Bodalia PN, Bryden PA, Davies PA, Lopez-Lopez JA, Okoli GN, Thom HZ, Caldwell DM, Dias S, Eaton D, Higgins JPT, Hollingworth W, Salisbury C, Savovic J, Sofat R, Stephens-Boal A, Welton NJ, Hingorani AD.

School of Social and Community Medicine, University of Bristol, Bristol, UK. jonathan.sterne@bristol.ac.uk

Department of Pharmacy, University College London Hospitals NHS Foundation Trust, London, UK. pritesh.bodalia@uclh.nhs.uk

OBJECTIVE: Compare the effectiveness and cost-effectiveness of all available directly acting oral anticoagulants (DOACs) for patients with atrial fibrillation (AF).

DESIGN: Systematic review, network meta-analysis (NMA), and cost-effectiveness analysis. We ranked the effects of warfarin, antiplatelet drugs and licensed doses of NOACs for seven efficacy and safety outcomes studied in patients with AF.

DATA SOURCES: MEDLINE, Pre-MEDLINE, EMBASE, and the Cochrane Library.

ELIGIBILITY CRITERIA: Published randomised trials evaluating the use of a NOAC, vitamin K antagonist, or antiplatelet agent in individuals with AF for stroke prevention**.**

RESULTS: We analysed 23 randomised trials involving 94,656 individuals: 13 compared a NOAC with warfarin dosed to achieve a target INR of 2-3. Apixaban 5mg bd (OR 0.79; 95% CI 0.66-0.94), dabigatran 150mg bd (OR 0.65; 0.52-0.81), edoxaban 60mg od (OR 0.86; 0.74-1.01) and rivaroxaban 20mg od (0.88; 0.74-1.03) reduced risk of stroke or systemic embolism compared with warfarin. Risk of stroke or systemic embolism was higher with edoxaban 60mg od (1.33; 1.02-1.75) and rivaroxaban 20mg od (1.35; 1.03-1.78) than dabigatran 150mg bd. Risk of all-cause mortality was lower with all DOACs than warfarin. Apixaban 5mg bd (0.71; 0.61-0.81), dabigatran 110mg bd (0.80; 0.69-0.93), edoxaban 30mg od (0.46; 0.40-0.54) and edoxaban 60mg od (0.78; 0.69-0.90) reduced risk of major bleeding compared with warfarin. Risk of major bleeding was higher with dabigatran 150mg bd than apixaban 5mg bd (1.33; 1.09-1.62), and with rivaroxaban 20mg bd than apixaban 5mg bd (1.45; 1.19-1.78) and edoxaban 60mg od (1.31; 1.07-1.59). Risk of intracranial bleeding was substantially lower for most DOACS compared with warfarin. Apixaban 5mg bd was ranked as the best intervention for most outcomes, and was cost-effective compared with warfarin.

CONCLUSION: Our NMA informs choice of DOAC for stroke prevention in AF. At licensed doses, a number of DOACs are of net benefit compared with warfarin, with apixaban ranked as the best intervention across a range of outcomes. A trial directly comparing DOACs would overcome the need for indirect comparisons to be made through NMA.